MICROCOPY RESOLUTION TEST CHART
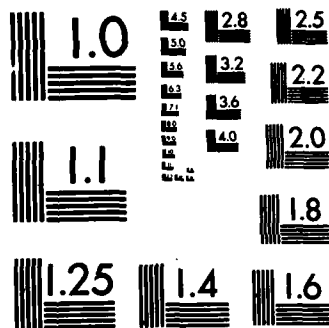
NATIONAL BUREAU OF STANDARDS-1963-A

Report Number 8

ALTERNATIVES TO OPTIMAL DETECTION

Y.F. Huang and J.B. Thomas

INFORMATION SCIENCES AND SYSTEMS LABORATORY

Department of Electrical Engineering and Computer Science
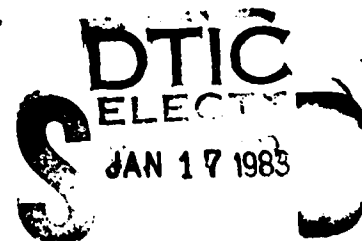Princeton University
Princeton, New Jersey 08544

NOVEMBER 1982

DTIC
ELECT
JAN 1 7 1983
E

83

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 8	2. GOVT ACCESSION NO. *A123 454*	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* ALTERNATIVES TO OPTIMAL DETECTION		5. TYPE OF REPORT & PERIOD COVERED Technical Report February '91–October '82
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Yih-Fang Huang and John B. Thomas		8. CONTRACT OR GRANT NUMBER(s) N00014-81-K-0146
9. PERFORMING ORGANIZATION NAME AND ADDRESS Information Science and Systems Lab. Dept. of Electrical Eng. & Computer Sci. Princeton Univ., Princeton, NJ 08544		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR SRO-013
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research (Code 411 SP) Department of the Navy Arlington, Virginia 22217		12. REPORT DATE November 1982
		13. NUMBER OF PAGES 124
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Non-Gaussian Noise	Skewed Noise
Edgeworth Series	Bounded Noise
Cornish-Fisher Expansion	

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

In this report, signal detection in non-Gaussian noise is
considered. This work can be divided roughly into two parts.
The essential purpose of the first part is to investigate the
existence and design of some sub-optimal detectors which are
modifications of the sample mean detector. The method used here
is based on orthogonal series expansions in terms of the Gauss-
ian distribution and the moments of the noise distribution.
(con't. on next pg)

DD, FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

(Abstract con't.)

The second part is concerned with signal detection in bounded noise. Both the known signal case and the unknown signal case are studied. An algorithm of a sequential nature is devised.

Chapters 2 and 3 are devoted to the first part of the work. To begin, an approximate locally optimal scheme, which can be considered as a modified linear detector, is discussed. This scheme evolves from a use of the Edgeworth series as a representation for the noise density function. Convergence of the approximate locally optimal statistic to the parent statistic is also discussed. Furthermore, the Cornish-Fisher inverse expansion is employed to investigate the effect of noise skewness on the sample mean detector and the sign detector. An asymptotic relation between the performance and the skewness measure is given. The sign detector is shown to be less sensitive than the sample mean detector to the noise skewness. A modified sample mean detector is devised.

Chapters 4 and 5 address signal detection in bounded noise. When both the signal and the bound on the noise magnitude are known constants, the sequential procedure will lead to a singular solution, if it terminates. A more practical way to consider this problem is to assume a random bound. Then a test which involves randomness in its performance measures (the false-alarm rate, the power and the error probability) will be encountered. When the signal is unknown, the problem is cast as an estimation problem. The estimate of the signal strength is used to decide whether the signal is present. A generalization of the problem is considered in Chapter 5 where a set-theoretic formulation is used.

ALTERNATIVES TO OPTIMAL DETECTION

by

Y.F. Huang and J.B. Thomas
Department of Electrical Engineering
and Computer Science
Princeton University
Princeton, N.J. 08544

ABSTRACT

In this report, signal detection in non-Gaussian noise is considered. This work can be divided roughly into two parts. The essential purpose of the first part is to investigate the existence and design of some sub-optimal detectors which are modifications of the sample mean detector. The method used here is based on orthogonal series expansions in terms of the Gaussian distribution and the moments of the noise distribution. The second part is concerned with signal detection in bounded noise. Both the known signal case and the unknown signal case are studied. An algorithm of a sequential nature is devised.

Chapters 2 and 3 are devoted to the first part of the work. To begin, an approximate locally optimal scheme, which can be considered as a modified linear detector, is discussed. This scheme evolves from a use of the Edgeworth series as a represent-ation for the noise density function. Convergence of the approx-imate locally optimal statistic to the parent statistic is also discussed. Furthermore, the Cornish-Fisher inverse expansion is employed to investigate the effect of noise skewness on the sample mean detector and the sign detector. An asymptotic re-lation between the performance and the skewness measure is given. The sign detector is shown to be less sensitive than the sample mean detector to the noise skewness. A modified sample mean detector is devised.

Chapter 4 and 5 address signal detection in bounded noise. When both the signal and the bound on the noise magnitude are known constants, the sequential procedure will lead to a singu-lar solution, if it terminates. A more practical way to consider this problem is to assume a random bound. Then a test which involves randomness in its performance measures (the false-alarm rate, the power and the error probability) will be encountered. When the signal is unknown, the problem is cast as an estimation prob-lem. The estimate of the signal strength is used to decide whether the signal is present. A generalization of this problem is considered in Chapter 5 where a set-theoretic formulation is used.

TABLE OF CONTENTS

Chapter 5 - A Set-Theoretic Approach to Detection Problems

Chapter 6 - Conclusions

CHAPTER 1

INTRODUCTION

1.1 Motivation

Extraction of signals from a noisy environment has been an important and much studied problem in statistical communication theory. In particular, the detection problem of deciding whether or not a signal is present by taking a sequence of noise contaminated observations has been of much interest to communication scientists and engineers. Various detection schemes have been discussed in a broad body of literature. In classical detection theory, parametric detection was of much concern due to the well developed theory of parametric statistical inference [1]. The Neyman-Pearson optimal detector and the locally optimal detector are two canonical examples. However, implementation of these (parametric) optimal schemes usually requires a fairly complete knowledge of the input model. In practice, this knowledge may not be available. Also, the performance may be sensitive to the inaccuracy of the presumed knowledge. Some attention thus has to be diverted to other alternatives.

Data from a number of natural environments show that noise is often non-Gaussian [2]. Nevertheless, due to the Central Limit Theorem, Gaussian noise is of much interest and, in fact, is the one for which the simplest solution can be obtained. In this case, it is well known that the sample mean detector is the Neyman-Pearson optimal detector [3]. Thus, in most practical situations, modifications of the sample mean

detector may yield a more satisfactory detector performance. In other words, these modifications could have "reasonably good" (or nearly optimal in some sense) performance for the presumed input model and could be less sensitive to model deviations. This motivates the first part of the report.

Motivation for the second part is due to the fact that a precise statistical model of the noise is difficult to obtain in many practical situations. Bounds on the magnitude of the noise, in contrast, may be more easily obtained. One may even argue that, in the real world, there exist no distributions with infinite support. Hence considerations of bounded noise should lead to more practical solutions. As a matter of fact, in control theory, bounded noise has been exploited in the literature for considering estimation problems [4], [5], [6].

1.2   Report Outline

Throughout this report   the noise processes are assumed to be discrete-time processes which may be obtained, for example, from the sampling of continuous-time processes. In addition, the detection problems considered here are detection of constant signals in additive noise. It should also be noted that all the chapters are complete by themselves and can be read in any order. The   report   is organized as follows:

Chapter 2 considers series expansions for the test statistics of two optimal detection schemes, the locally optimal statistic and the sample-mean statistic which is Neyman-Pearson optimal for the Gaussian case. For the former, the Edgeworth series is employed to obtain a general

representation of the noise density function. The locally optimal statistic is then evaluated via this representation. Since only a few lower order moments of the noise are required, the solutions are applicable to a wider class of noise distributions. The second part of this chapter utilizes the Cornish-Fisher inverse expansion for a sequence of i.i.d. random variables. The sensitivity of the sample mean detector to the underlying noise skewness can then be studied. An asymptotic relation between the performance (false-alarm, power, and probability of error) and the noise skewness measure is obtained.

The Cornish-Fisher expansion is used again in Chapter 3 for further investigation of the effect of noise skewness on detector performance. A modified sample mean detector is proposed whose performance is asymptotically indifferent to noise skewness. Simulation results are provided to verify the analysis. A preliminary case study of a natural noise environment is also given here.

The latter part of this report discusses the detection problem from a rather different point of view. Instead of noise statistics, bounds for the noise magnitude are assumed to be known in Chapter 4. To make the problem more practical, the bound is assumed to be a random variable with some known distribution. Then a sequential procedure is discussed. Detection of an unknown signal in bounded noise is also addressed in this chapter. The problem is cast as an estimation problem for the signal.

Chapter 5 is essentially concerned with a generalized version of the problem discussed in the second part of Chapter 4. A set-theoretic for-

mulation of the problem is used here to consider a "multi-channel" detection problem. The results presented in this chapter are somewhat preliminary due to the difficulty involved in performance evaluation. Chapter 6 concludes this report and discusses some possibilities for future research.

References

[1] H.L. Van Trees, Detection, Estimation, and Modulation Theory, Part I. New York: John Wiley, 1968.

[2] R.F. Dwyer, "Fram II Single Channel Ambient Noise Statistics," *101st Meeting of The Acoustical Society of America*, May, 1981.

[3] J. Marcum, "Statistical Theory of Target Detection by Puls Radar," *IRE Trans. Information Theory*, vol. IT-6, pp. 59-267, 19

[4] F.C. Schweppe, "Recursive State Estimation: Unknown b Bounded Errors and System Inputs," *IEEE Trans. Automat. Con.* ol. AC-13, pp.22-28, 1968.

[5] D.P. Bertsekas and I.B. Rhodes, "Recursive State Estimation for a Set-Membership Description of Uncertainty," *IEEE Trans. on Automat. Contr.*, Vol. AC-16, pp. 117-128, 1971.

[6] M. Milanese and G. Belforte, "Estimation Theory and Uncertainty Intervals Evaluation in Presence of Unknown But Bounded Errors: Linear Families of Models and Estimators," *IEEE Trans. on Automat. Contr.*, Vol. AC-27, pp. 408-414, 1982.

CHAPTER 2

ON THE APPLICATION OF SERIES EXPANSIONS TO DETECTION PROBLEMS

2.1 Introduction

The problem of using series expansions to represent an unknown function in terms of a known function has been of much interest to scientists and engineers. In the literature of statistics, seeking for a general representation of probability distribution functions or density functions has been an extensively studied problem, [1] - [4]. A representation based on series expansions in terms of the moments is particularly relevant when sequences of independent and identically distributed (i.i.d.) random variables are considered. One of the most common series expansions used to represent an unknown probability density function by a known density function is the Gram-Charlier series. The Gram-Charlier series is an expansion in a series of orthogonal polynomials, i.e., the Tchebycheff-Hermite polynomials, which are derived from the normal density function. Unfortunately, straightforward applications of the Gram-Charlier series do not usually lead to satisfactory solutions. It is well known in the literature that regrouping of the Gram-Charlier series, e.g. the Edgeworth series [5],[6], will yield better results; although generalization of the Gram-Charlier series is also a possible way for improving the result [7].

In the context of statistical communications, the method of series expansions has also been used to evaluate approximately the probability of errors in radar detection [8], and to estimate error probabilities in

digital communication systems [9] - [12]. This method is employed in this chapter to consider detection problems, although for somewhat different usages. To begin with, it is used here to obtain a general representation for the noise density function. Then the locally optimum detection scheme is devised based on this representation. The essential purpose of this approach is to reduce the effect of small model deviations on the locally optimal detectors. This is because series representation only requires a knowledge of the moments of the noise distribution. Thus when a finite number of terms of the series are used, the result can be expected to hold approximately for a certain family of noise distributions.

The other utility of the series expansions in this chapter is the application of the Cornish-Fisher inverse expansion [13] for a random variable in terms of its moments. This expansion is used to obtain a representation for a test statistic, the sample mean of a sequence of i.i.d. random variables. In this way, a study relating the performance of the sample mean detector to the underlying noise skewness can be facilitated.

This chapter is organized as follows : the next section discusses derivations of the Tchebycheff-Hermite polynomials and the Gram-Charlier series. Section 2.3 addresses some difficulties associated with the Gram-Charlier series and presents a regrouped version of the Gram-Charlier series, namely, the Edgeworth series. Section 2.4 considers a detection scheme using an approximate locally optimal statistic which is obtained by using the Edgeworth series representation for the noise density function. Section 2.5 utilizes the Cornish-Fisher inversion series for

the sample mean of a sequence of i.i.d. random variables to study the effect of noise skewness on the sample mean detector. Section 2.6 concludes this chapter.

2.2 Tchebycheff-Hermite Polynomials and Gram-Charlies Series

Derivations of the Gram-Charlier series can be found in many standard statistics texts (see e.g. [5],[6]). For completeness of discussion, one which follows mostly from [6] is given here. The derivation will start with the definition of the Tchebycheff-Hermite polynomials. Consider a Gaussian density function $f_0(x)$ given by

$$f_0(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} \tag{1}$$

Taking successive derivatives of $f_0(x)$ with respect to x yields

$$f'_0(x) = -\frac{x}{\sigma^3} f_0(x)$$

$$f''_0(x) = (\frac{x^2}{\sigma^4} - 1) f_0(x)$$

$$f_0^{(3)}(x) = (-\frac{x^3}{\sigma^6} + \frac{3x}{\sigma^4}) f_0(x)$$

$$f_0^{(4)}(x) = (\frac{x^4}{\sigma^8} - 6\frac{x^2}{\sigma^6} + 3) f_0(x)$$

$$f_0^{(5)}(x) = (-\frac{x^5}{\sigma^{10}} + 10\frac{x^3}{\sigma^8} - 15\frac{x}{\sigma^6}) f_0(x)$$

etc. The Tchebycheff-Hermite polynomial is defined by the following equation

$$f_0^{(n)}(x) = (-1)^n \sigma^{-n} H_n(x) f_0(x)$$

Thus

$$H_0(x) = 1$$
$$H_1(x) = \frac{x}{\sigma}$$

$$H_2(x) = \frac{x^2}{\sigma^2} - 1$$

$$H_3(x) = \frac{x^3}{\sigma^3} - 3\frac{x}{\sigma}$$

$$H_4(x) = \frac{x^4}{\sigma^4} - 6\frac{x^2}{\sigma^2} + 3$$

$$H_5(x) = \frac{x^5}{\sigma^5} - 10\frac{x^3}{\sigma^3} + 15\frac{x}{\sigma}$$

and so on. Now, from Eq. (1),

$$f_0(x-t) = f_0(x) e^{\frac{tx}{\sigma^2} - \frac{t^2}{2\sigma^2}}$$

then, by Taylor's series expansion, we have

$$f_0(x-t) = \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} t^i f_0^{(i)}(x)$$

$$= \sum_{i=0}^{\infty} \frac{t^i}{i!} H_i(x) f_0(x)$$

Thus, a general formula for the Tchebycheff-Hermite polynomials is

$$H_n(x) = \frac{x^n}{\sigma^n} - \frac{n^{[2]}}{2 \cdot 1!} \frac{x^{n-2}}{\sigma^{n-2}} + \frac{n^{[4]}}{2^2 \cdot 2!} \frac{x^{n-4}}{\sigma^{n-4}} - \frac{n^{[6]}}{2^3 \cdot 3!} \frac{x^{n-6}}{\sigma^{n-6}} + \ldots \quad (2)$$

where

$$n^{[r]} \triangleq n(n-1)(n-2)\ldots(n-r+1) \qquad r \leq n$$

Apparently, the polynomials are of degree n in x with the coefficient of x^n being unity. Also, they are orthogonal over the interval $(-\infty, \infty)$ with respect to the weighting function $f_0(x)$, namely

$$\int_{-\infty}^{\infty} f_0(x) H_m(x) H_n(x) dx = \delta_{mn} n! \qquad (3)$$

where

$$\delta_{mn} \begin{cases} = 0 & \text{if } m \neq n \\ = 1 & \text{if } m = n \end{cases}$$

is the Kronecker delta function. Furthermore, successive Tchebycheff-Hermite polynomials satisfy the following relations :

$$\frac{d}{dx}H_n(x) = \frac{n}{\sigma}H_{n-1}(x) \tag{4}$$

$$H_n(x) - \frac{x}{\sigma}H_{n-1}(x) + (n-1)H_{n-2}(x) = 0 \tag{5}$$

Now, the probability density function $f(x)$ of a random variable (with zero mean and known variance σ^2) can be represented in terms of the Tchebycheff-Hermite polynomials, i. e.

$$f(x) = \sum_{i=0}^{\infty} c_i H_i(x) f_0(x) \tag{6}$$

This is the so-called Gram-Charlier series. Due to the orthogonal property of the polynomials, Eq. (3), the coefficients c_r in Eq. (6) can be determined by multiplying both sides of Eq. (6) by $H_r(x)$ and integrating the products from $-\infty$ to ∞. Hence

$$c_r = \frac{1}{r!} \int_{-\infty}^{\infty} f(x)H_r(x)dx \tag{7}$$

Thus

$$c_0 = 1$$
$$c_1 = 0$$
$$c_2 = 0$$
$$c_3 = \frac{m_3}{6\sigma^3}$$
$$c_4 = \frac{1}{24}(\frac{m_4}{\sigma^4} - 3)$$
$$c_5 = \frac{1}{120}(\frac{m_5}{\sigma^5} - 10\frac{m_3}{\sigma^3}) \ .$$

etc., where the m_i's are the i-th central moments of $f(x)$. In some situations, the coefficients c_r may be better represented in terms of the cumulants. Thus,

$$c_3 = \frac{\kappa_3}{6\kappa_2^{3/2}}$$

$$c_4 = \frac{\kappa_4}{24\kappa_2^2} \tag{8}$$

$$c_5 = \frac{\kappa_5}{120\kappa_2^{5/2}}$$

etc.. Then the series expansion for $f(x)$ is given by

$$f(x) = f_0(x)[\ 1 + \frac{1}{6}\frac{m_3}{\sigma^3}H_3(x) + \frac{1}{24}(\frac{m_4}{\sigma^4} - 3)\ H_4(x) + \ldots\ ]$$

$$= f_0(x)[\ 1 + \frac{1}{6}\frac{\kappa_3}{\kappa_2^{3/2}}H_3(x) + \frac{1}{24}\frac{\kappa_4}{\kappa_2^2}H_4(x) +$$

$$\frac{1}{120}\frac{\kappa_5}{\kappa_2^{5/2}}H_5(x) + \ldots\ ] \qquad (9)$$

It is known that a sufficient condition for this series to be convergent (at every continuity point of $f(x)$) is [5]

$$\int_{-\infty}^{\infty} e^{x^2/4}f(x)dx < \infty \qquad (10)$$

Obviously, an infinite series is somewhat impractical. It has to be truncated at some point. However, truncation of the series will result in some difficulties. The series may not converge fast enough so that the truncated series will be a satisfactory representation of $f(x)$. Furthermore, the order of magnitude of successive terms of the series is not necessarily monotone decreasing. These problems will be discussed in more detail in the next section.

2.3 The Edgeworth Series

As discussed in the previous section, the Gram-Charlier series, Eq. (9), does not necessarily provide a good approximation to the probability density function. Generally speaking, a truncated version of the Gram-Charlier series can be a good approximation only when the parent probability density function is nearly Gaussian. A well known problem is that the order of magnitude of the terms in the Gram-Charlier series is not

steadily decreasing. To be more specific, suppose, for example, that one wants to represent the probability density function of the sample mean of a sequence of i.i.d. random variables $\{X_i\}$ in terms of the Gaussian density function. Without loss of generality, it is assumed here that the distribution of $\{X_i\}$ has zero mean and unity variance. Now, let

$$S_M = \frac{1}{M}\sum_{i=1}^{M} X_i$$

From Eq. (9), the p. d. f. of S_M can be represented as follows

$$f_S(s) = f_0(s) \left[1 + \frac{1}{6}\kappa'_3 H_3(s) + \frac{1}{24}\kappa'_4 H_4(s) + \frac{1}{120}\kappa'_5 H_5(s) \right.$$
$$\left. + \frac{1}{720}(\kappa'_6 + 10\kappa'_3{}^2)H_6(s) + \ldots \right] \tag{11}$$

where

$$\kappa'_3 = \frac{\kappa_3}{\sqrt{M}}$$

$$\kappa'_4 = \frac{\kappa_4}{M}$$

$$\kappa'_5 = \frac{\kappa_5}{M^{3/2}}$$

$$\kappa'_6 = \frac{\kappa_6}{M^2}$$

and the κ'_i are the cumulants of S_M while the κ_i are the cumulants of X_i. In fact, if Eq. (11) is to be written as follows

$$f_S(s) = \sum_{r=0}^{\infty} f_0(s)\, c_r\, H_r(s) \tag{12}$$

then a general formula for c_r is [5]

$$c_r = \frac{a_{r1}M + a_{r2}M^2 + \cdots + a_{r[r/3]}M^{[r/3]}}{r!\, M^{r/2}} \tag{13}$$

where $[r/3]$ denotes the greatest integer $\leq r/3$, and the a_{ri} are polynomials in κ_r. The a_{ri} are independent of M. Also, as M tends to infinity,

$$c_r = O \left(M^{[r/3]-r/2} \right) \tag{14}$$

Thus, the order of magnitude of the terms of the Gram-Charlier series is clearly not steadily decreasing as r increases. The following table illustrates the order of magnitude in terms of powers of M that the coefficients c_r involve.

Order of c_r	r
$M^{-1/2}$	3
M^{-1}	4,6
$M^{-3/2}$	5,7,9
M^{-2}	8,10,12
$M^{-5/2}$	11,13,15

Now, suppose that a partial sum of the series, Eq. (12), is to be calculated such that all terms of magnitude order less than M^{-1} are truncated. Then, it can be seen from the table that all terms up to $r = 6$ should be considered. Thus, the moments up to sixth order will be needed. However, a careful examination of Eqs. (12) and (13) reveals that the moments of order higher than the fourth really do not involve terms of magnitude order greater than M^{-1}. Therefore, the requirement of the moments up to sixth order is unnecessary if all terms of order less than M^{-1} are to be neglected. Similarly, if one proceeds further to include terms containing the factors $M^{-3/2}$, M^{-2}, etc., the same redundancy will be encountered.

From the above discussion, it is clear that a regrouping of terms in the Gram-Charlier series will improve the efficiency of the approximation. The Edgeworth series, which is a regrouped version of the Gram-Charlier series, actually provides more satisfactory solutions. A general form of

the Edgeworth series expansion for an unknown density function in terms of its moments and the normal density function is given in the following

$$f(x) = f_0(x) [1 + c_3 H_3(x) + c_4 H_4(x) + c_6 H_6(x) +$$
$$c_5 H_5(x) + c_7 H_7(x) + c_9 H_9(x) + \ldots] \qquad (15)$$

where the c_r's are given in Eq.(8). As can be seen from Eq. (15), the order of terms in the Edgeworth series follows directly those given in the table of the last page. One can see that this series gives a straightforward expansion in powers of $M^{-1/2}$. Also, calculation of the terms in the Edgeworth series up to a certain order of magnitude does not require a knowledge of any moments or cumulants that are not really necessary. Thus the redundancy involved in the Gram-Charlier series discussed previously is eliminated. Furthermore, terms of the Edgeworth series should be taken by groups of the same magnitude order; thus any partial sum of this series is an asymptotic expansion of the parent density function in powers of $M^{-1/2}$, with a remainder term of the same order as the first term neglected. More complete discussions on the derivation of the Edgeworth series can be found in [5],[6].

2.4 Edgeworth Series Expansion for Locally Optimum Statistics

In detection problems, optimal detection schemes usually require a large amount of knowledge concerning the noise environment. Also, it is not unusual that these optimal schemes are complicated and not easy to implement. Furthermore, the performance may be sensitive to the accuracy of the presumed noise statistics. Thus alternatives that involve less complexity and less sensitivity are always desirable. The locally optimal

detectors are known to have maximum rate of increase in the probability of detection when the signal strength is equal to zero. Hence, they are useful schemes when the signal strength is small. However, their implementation require a complete knowledge of the underlying noise density function and the test statistics become complicated or even intractable very easily. In applications where the noise density function is only known to be nearly Gaussian (or, specifically, it satisfies Eq. (10)), the Edgeworth series expansion discussed in the previous section may be used as a general representation of the noise density function. Then the locally optimum test statistics can be obtained based on this representation of the density function. The result thus obtained is expected to be less sensitive to the deviations of noise statistics as only the lower-order moments of the noise distribution are required to obtain this representation.

Suppose that the noise distribution has zero mean and unity variance (a generalization of the following discussion with assumption of an arbitrary variance is trivial). Now consider a series representation, $g(x)$, up to a second order approximation of the noise density function $f(x)$ such that

$$| f(x) - g(x) | < \frac{C}{N} \qquad \text{for almost all } x \in (-\infty, \infty) ,$$

where C is a constant. Then, from Eq. (15), the noise density function can be represented by

$$g(x) = f_0(x)[1 + c_3 H_3(x) + c_4 H_4(x) + c_6 H_6(x)] \qquad (16)$$

where

$$f_0(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

is the standard normal density function. Now the derivative of $g(x)$ is equal to

$$g'(x) = f'_0(x) [1 + c_3 H_3(x) + c_4 H_4(x) + c_6 H_6(x)] + \qquad (17)$$
$$f_0(x) [3c_3 H_2(x) + 4c_4 H_3(x) + 6c_6 H_5(x)]$$

If it is further assumed that the noise density function is symmetric with respect to the origin, then $c_3 = 0$ and Eqs. (16) and (17) become

$$g(x) = f_0(x) [1 + c_4 H_4(x) + c_6 H_6(x)] \qquad (18)$$

and

$$g'(x) = f'_0(x) [1 + c_4 H_4(x) + c_6 H_6(x)] + \qquad (19)$$
$$f_0(x) [4c_4 H_3(x) + 6c_6 H_5(x)]$$

Now the locally optimal statistic can be approximated by

$$T(x) = - \frac{g'(x)}{g(x)} = x - \frac{4c_4 H_3(x) + 6c_6 H_5(x)}{1 + c_4 H_4(x) + c_6 H_6(x)} \qquad (20)$$

The second term in Eq. (20) can be regarded as a correction term due to the deviation from the Gaussian model. The performance of this non-linear detection scheme is evaluated here via the asymptotic relative efficiency (ARE) with respect to the linear detector. The ARE of this approximated locally optimal detector comparing to the linear detector is, [14],

$$ARE_{alod,ld} = \frac{[\int_{-\infty}^{\infty} (\frac{\partial}{\partial x} T(x)) f(x) dx]^2}{Var[T(X)]} \qquad (21)$$

Since $T(x)$ is an odd function in x, Eq. (20), and it is assumed that $f(x)$ is symmetric with respect to the origin, thus $E[T(x)] = 0$. Furthermore, if one assumes that

$$\lim_{|x| \to \infty} x^n f(x) = 0 \qquad \text{for any integer } n, \qquad (22)$$

then

$$ARE_{alod,ld} = \frac{[\int_{-\infty}^{\infty} T(x) f'(x)dx]^2}{Var[T(X)]}$$

On the other hand, it can be shown easily that the *ARE* of the locally optimal detector as compared to the linear detector is given by

$$ARE_{lod,ld} = \int_{-\infty}^{\infty} \frac{f'^2(x)}{f(x)} dx .$$

if Eq. (22) holds.

From Eq. (20), it is trivial to see that

$$\sigma_T^2 \triangleq Var[T(X)] = 1 - 2E[Xh(X)] + E[h^2(X)]$$

where

$$h(X) = \frac{4c_4 H_3(x) + 6c_6 H_5(x)}{1 + c_4 H_4(x) + c_6 H_6(x)}$$

By the Central Limit Theorem, it can be shown that (Appendix 2.A)

$$ARE_{alod,ld} = \frac{1}{\sigma_T^2},\tag{23}$$

therefore,

$$ARE > 1 \iff \sigma_T < 1 \iff E[h^2(X)] < 2E[Xh(X)]$$

Note that

$$E\{h^2(X)\} = \int_{-\infty}^{\infty} h^2(x)f(x)dx = 2\int_0^{\infty} h^2(x)f(x)dx$$

and

$$E\{Xh(X)\} = \int_{-\infty}^{\infty} xh(x)f(x)dx = 2\int_0^{\infty} xh(x)f(x)dx$$

also, $f(x)$ is a p.d.f. and is thus non-negative for $x \in (-\infty,\infty)$. Hence a sufficient condition for the $ARE_{alod,ld}$ to be greater than unity is

$$2xh(x) > h^2(x) \qquad \forall \, x \in (0, \infty)\tag{24}$$

However, it is not necessarily true that Eq. (24) holds. Detailed discussions on Eq. (24) are given in Appendix 2.B. It can be expected ,nevertheless, that when the noise model is nearly Gaussian, the $ARE_{alod,ld}$ is

greater than unity.

Finally, it should be noted that, in general, it is not necessarily true that $g'(x)$, Eq. (17), will converge to $f'(x)$ even if $g(x)$ converges to $f(x)$. A sufficient condition to be satisfied is the Tauberian condition [15] which requires that $f'(x)$ be monotone nondecreasing. Unfortunately, for some density functions, this condition may not be satisfied. However, practically speaking, convergence is an asymptotic property which does not usually provide enough information as to how well the series expansion approximates the parent function, especially when only a finite number (or even a small number) of terms are used. In other words, an essential issue is whether or not the expansion yields a good approximation to the parent function. Moreover, in detection problems, it is of more concern that the resulting detection scheme leads to a good performance measure, e.g. an *ARE* greater than unity. Figs. 2.A and 2.B show curves describing the $T(x)$ given in Eq. (19), together with the test statistics of the locally optimum and the linear detectors using mixture noise models. These models are Gaussian slightly contaminated with Laplace noise. The $ARE_{alod,ld}$ for these two examples are greater than one; although the $T(x)$ is a good approximation to the locally optimal statistic only in the neighborhood of the origin. This result provides some positive illustration of applications of the method discussed here.

2.5 Cornish-Fisher Expansion for Test Statistics

In this section, the Cornish-Fisher expansion of a random variable in terms of its moments, [13], will be employed for the sample mean of a

Fig. 2.A Locally Optimal Statistics For Mixture Model

$$\gamma^2 \triangleq \frac{\sigma_2^2}{\sigma_1^2}$$

$$f(x) = (1-\varepsilon)\frac{1}{\sqrt{2\pi}\sigma_1}e^{-x^2/2\sigma_1^2} + \varepsilon\,\frac{\alpha}{2}e^{-\alpha|x|}$$

$$\sigma_2^2 = \frac{2}{\alpha^2}$$

$$\varepsilon = 0.05 \quad \gamma = 0.8$$

$$ARE_{alod,ld} = 1.001$$
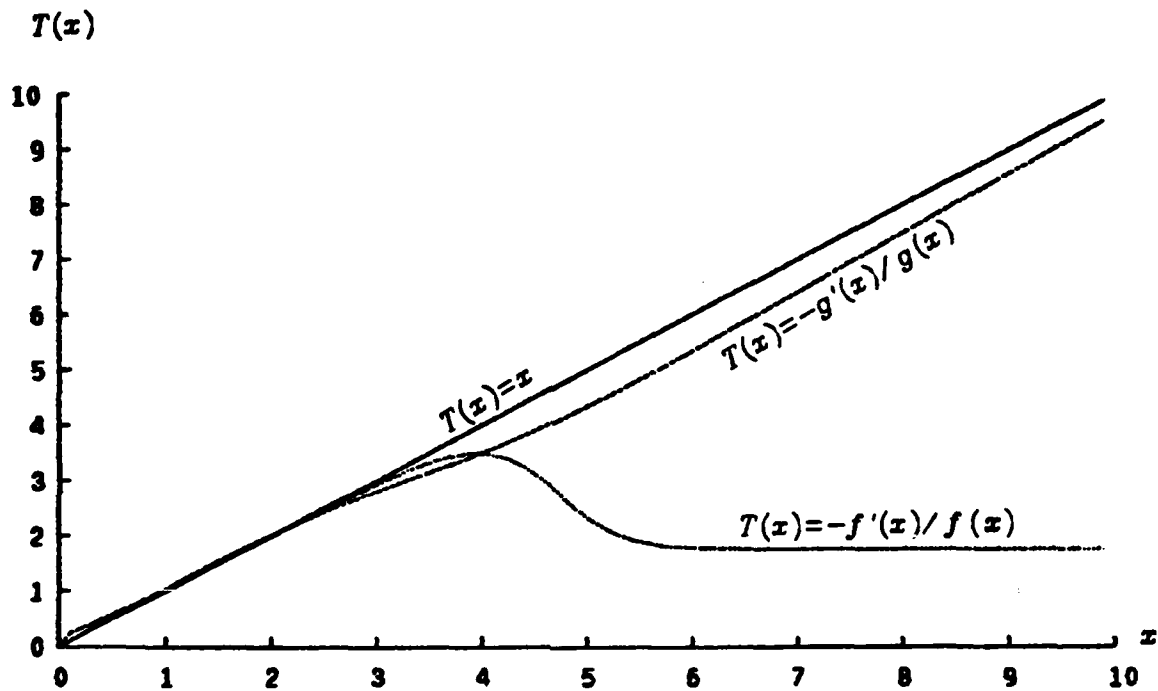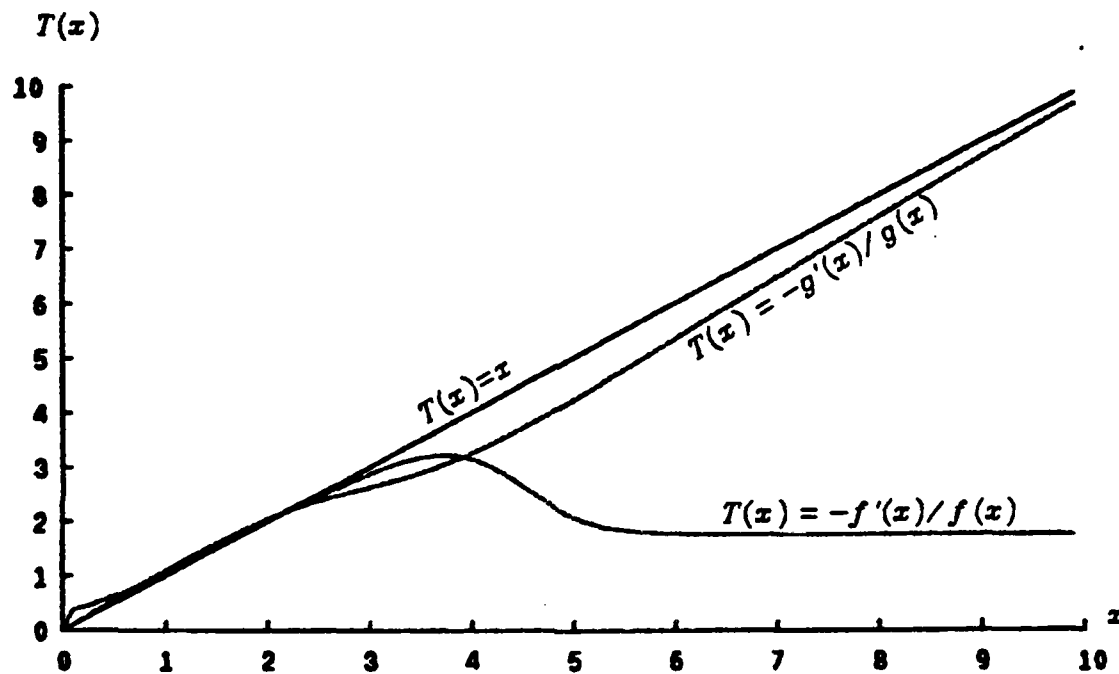
$$ARE_{lod,ld} = 1.008$$

Fig. 2.B Locally Optimal Statistics For Mixture Model

$$\gamma^2 \triangleq \frac{\sigma_2^2}{\sigma_1^2}$$

$$f(x) = (1-\varepsilon) \frac{1}{\sqrt{2\pi}\sigma_1} e^{-x^2/2\sigma_1^2} + \varepsilon \frac{\alpha}{2} e^{-\alpha|x|}$$

$$\sigma_2^2 = \frac{2}{\alpha^2}$$

$$\varepsilon = 0.1 \quad \gamma = 0.8$$

$$ARE_{alod,ld} = 1.003$$

$$ARE_{lod,ld} = 1.02$$

sequence of i.i.d. random variables. In detection problems, this sample mean is known to be the test statistic of the sample mean detector which is the Neyman-Pearson optimal detector when the underlying noise is Gaussian. However, when the noise density is not Gaussian due to skewness, the performance of the linear detector may deteriorate. Little work has been done on relating performance of the sample mean detector to noise skewness, partly because the conventional skewness measure [6] does not appear explicitly in the likelihood ratio or the test statistics. The objective here is to use a series expansion for the test statistic in terms of its moments such that an explicit relation between the noise skewness and the performance of the detector can be obtained. In order to facilitate the study, the underlying noise will be assumed to be only slightly skewed and nearly Gaussian. This assumption enables us to use the first few terms of the Cornish- Fisher expansion and obtain a good approximation to the test statistic.

Now the following canonical binary hypothesis testing problem will be considered:

$$H_0 : X_i = N_i$$
$$H_1 : X_i = N_i + s, \quad s > 0 \quad i = 1, 2, \ldots, M$$

It is assumed here that, under H_0, the p. d. f. of X_i is $f(x_i)$ with zero mean, known variance σ^2 and known third-order central moment μ_3, where μ_3 is assumed to be small. Furthermore, the observation sequence $\{X_i\}$ is supposed to be independent and identically distributed. Now, consider the following test statistic

$$T(X) = \bar{X} \triangleq \frac{1}{M} \sum_{i=1}^{M} X_i$$

This $T(X)$ can be written in terms of the Cornish-Fisher expansion, namely,

under H_0: $T(X) = \dfrac{\sigma}{\sqrt{M}} Z + \dfrac{\mu_3}{6\sigma^2 M}(Z^2 - 1) + O(M^{-3/2})$

and (25)

under H_1: $T(X) = \dfrac{\sigma}{\sqrt{M}} Z + s + \dfrac{\mu_3}{6\sigma^2 M}(Z^2 - 1) + O(M^{-3/2})$

where Z is a standard normal random variable and $O(M^{-3/2})$ stands for the terms which converge to zero as fast as $M^{-3/2}$ when M tends to infinity.

It is clear that, in order to make the problem tractable, the series in Eq. (25) has to be truncated. Since it has been assumed that μ_3 is small, the truncated series will be a good approximation to $T(X)$. Now, define two random sequences $\{Y_0(M)\}$ and $\{Y_1(M)\}$ by

$$Y_0(M) = \dfrac{\sigma}{\sqrt{M}} Z + \dfrac{\mu_3}{6\sigma^2 M}(Z^2 - 1)$$

and (26)

$$Y_1(M) = \dfrac{\sigma}{\sqrt{M}} Z + s + \dfrac{\mu_3}{6\sigma^2 M}(Z^2 - 1)$$

Then, it can be shown that $\{Y_0(M)\}$ and $\{Y_1(M)\}$ converge to $T(X)$ almost surely under H_0 and H_1 respectively. Hence the performance of the detector which uses $T(X)$ as the test statistic can be evaluated approximately via the statistics of $\{Y_0(M)\}$ and $\{Y_1(M)\}$, provided that M is sufficiently large.

Let $A = \mu_3/6\sigma^2 M$ and $B = \sigma/\sqrt{M}$, then Eq. (26) becomes $Y_0(M) = AZ^2 + BZ - A$ and $Y_1(M) = AZ^2 + BZ - A + s$. Note that Z is a standard normal random variable. Thus the probability density functions of Y_0 and Y_1 are given by

$$f_{Y_0}(y) \begin{cases} = \dfrac{1}{\sqrt{2\pi}} Q^{-1}(y)\{\Phi(-\dfrac{B}{2A})\exp[-\dfrac{1}{2}(\dfrac{Q(y)+B}{2A})^2] + \\ \qquad \Phi(\dfrac{B}{2A})\exp[-\dfrac{1}{2}(\dfrac{Q(y)-B}{2A})^2] \} & \text{if } y \ge -A-\dfrac{B^2}{4A} \\ = 0 & \text{otherwise} \end{cases} \qquad (27)$$

where

$$Q(y) = \sqrt{B^2 + 4A^2 + 4Ay} \qquad (28)$$

and

$$f_{Y_1}(y) = f_{Y_0}(y-s)$$

The false-alarm rate α is given by

$$\alpha = \int_{T_0}^{\infty} f_{Y_0}(y)dy \qquad (29)$$

where T_0 is the threshold. From Eq. (28), it follows that

$$dQ(y) = 2A \; Q^{-1}(y) \; dy \qquad (30)$$

Substituting Eqs. (27) and (30) into Eq. (29) yields

$$\alpha = \int_{Q(T_0)}^{\infty} \Phi(-\dfrac{B}{2A}) \dfrac{1}{2A\sqrt{2\pi}} \exp[-\dfrac{1}{2}(\dfrac{Q+B}{2A})^2]dQ + $$
$$\int_{Q(T_0)}^{\infty} \Phi(\dfrac{B}{2A}) \dfrac{1}{2A\sqrt{2\pi}} \exp[-\dfrac{1}{2}(\dfrac{Q-B}{2A})^2]dQ$$

Thus

$$\alpha = \Phi(-\dfrac{B}{2A})\Phi(-\dfrac{T_Q+B}{2A}) + \Phi(\dfrac{B}{2A})\Phi(-\dfrac{T_Q-B}{2A}) \qquad (31)$$

where

$$T_Q \triangleq Q(T_0)$$

Similarly, the power β is given by

$$\beta = \Phi(-\dfrac{B}{2A})\Phi(-\dfrac{Q(T_0-s)-B}{2A}) \qquad (32)$$

If the skewness is defined as

$$\xi = \dfrac{\mu_3}{\sigma^3}$$

then, it is obvious that both α and β are dependent upon the noise

skewness, ξ, via the variables $\frac{B}{2A}$, $\frac{T_Q}{2A}$ and $\frac{Q(T_0-s)}{2A}$. In particular,

$$\frac{T_Q \pm B}{2A} = \sqrt{1 + \frac{9M}{\xi^2} + \frac{6MT_0}{\sigma\xi}} \pm 3\frac{\sqrt{M}}{\xi}$$

Fig. 2.C gives a set of curves which depicts the normalized rate of change in false-alarm rate, power, and probability of error as functions of the noise skewness. It is assumed that the underlying noise density, which is described by Eq. (27), is positively skewed. One can see that both the false-alarm rate and the probability of error are monotone increasing functions of the noise skewness ξ. On the other hand, the power stays nearly unchanged. This illustrates the performance deterioration of the sample mean detector when the underlying noise is skewed. A density function, Eq. (27), with a skewness measure $\xi = 0.6$ is shown in Fig. 2.D together with the corresponding Gaussian density function. For the initial false-alarm rate $\alpha = 10^{-5}$, this noise density function corresponds to a 100% increase in false-alarm rate. Although these two density functions are hardly distinguishable, from Fig. 2.D, the difference between their associated false-alarm rates is rather significant. Fig. 2.E plots the Neyman-Pearson optimal statistics associated with these two noise models given in Fig. 2.D.

In Eq. (27), by use of the L'Hospital's rule, it can be shown that

$$\lim_{\xi \to 0} f_{Y_0}(y) = g(y)$$

where

$$g(y) = \frac{1}{\sqrt{2\pi B^2}} \exp\left[-\frac{1}{2}\left(\frac{y}{B}\right)^2\right] \quad \text{for} \quad -\infty < y < \infty$$
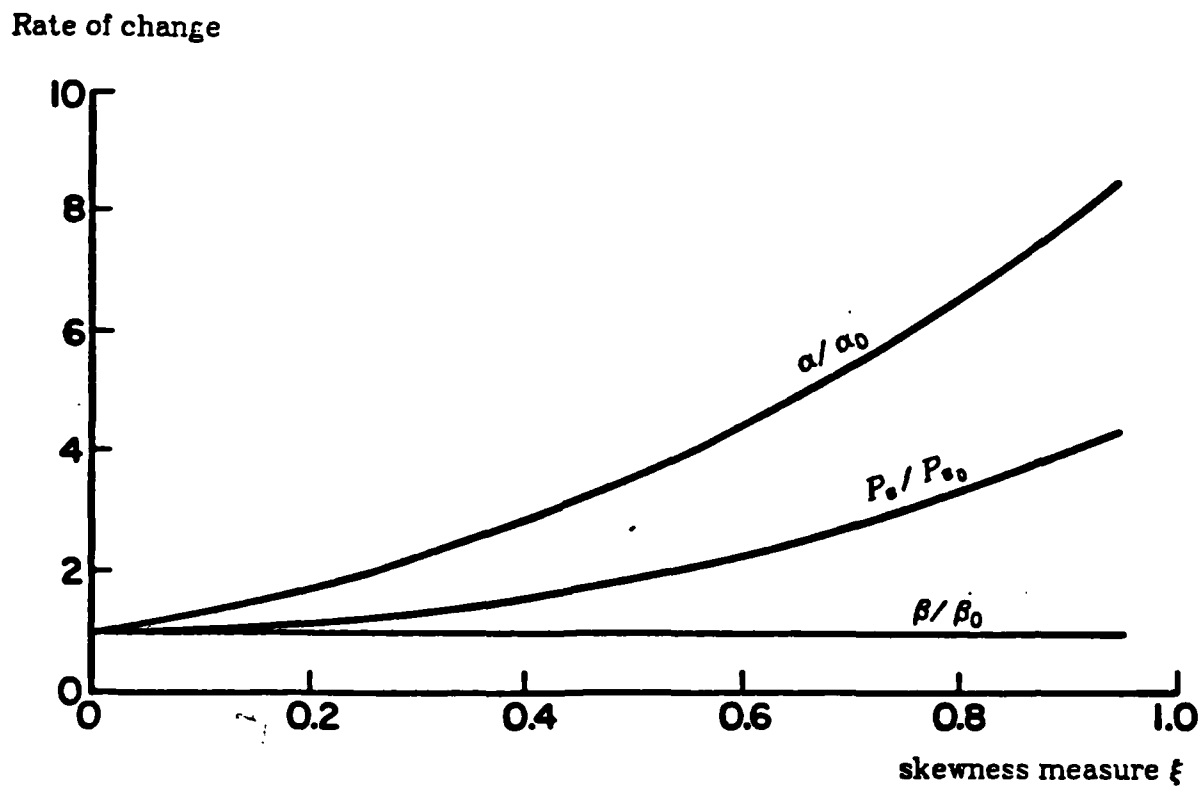
and thus

Fig. 2.C Rate of Change In False-alarm Rate, Power,

and Probability Of Error

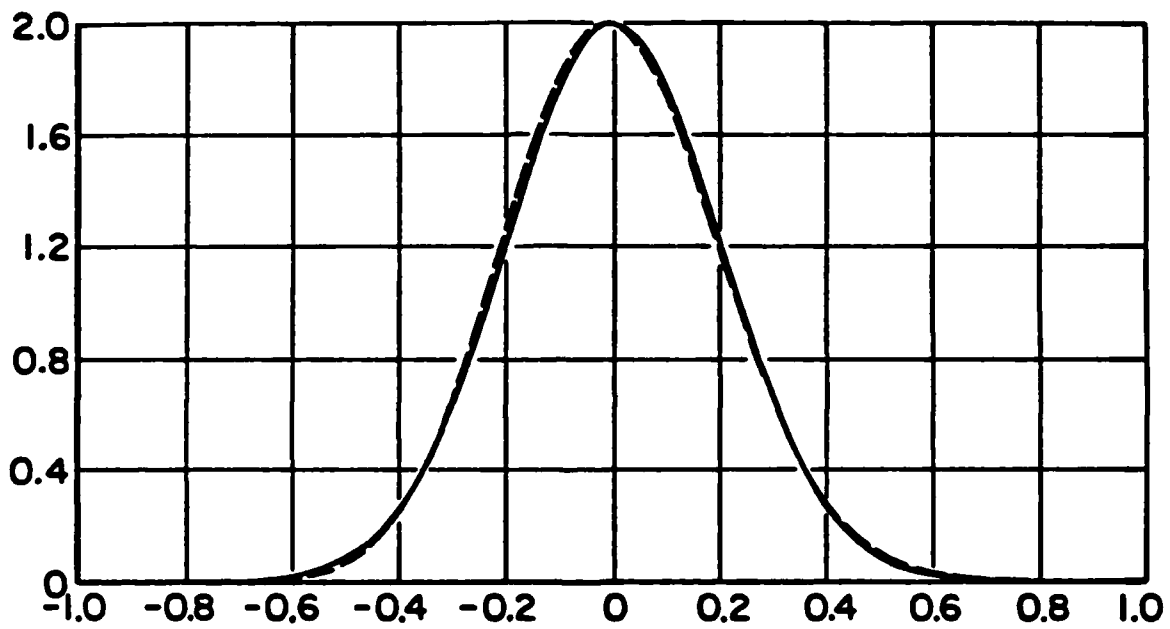Sample Size $M = 100$

$\alpha_0 = 10^{-10}$

Fig. 2.D Probability Density Functions of Eq. (27)
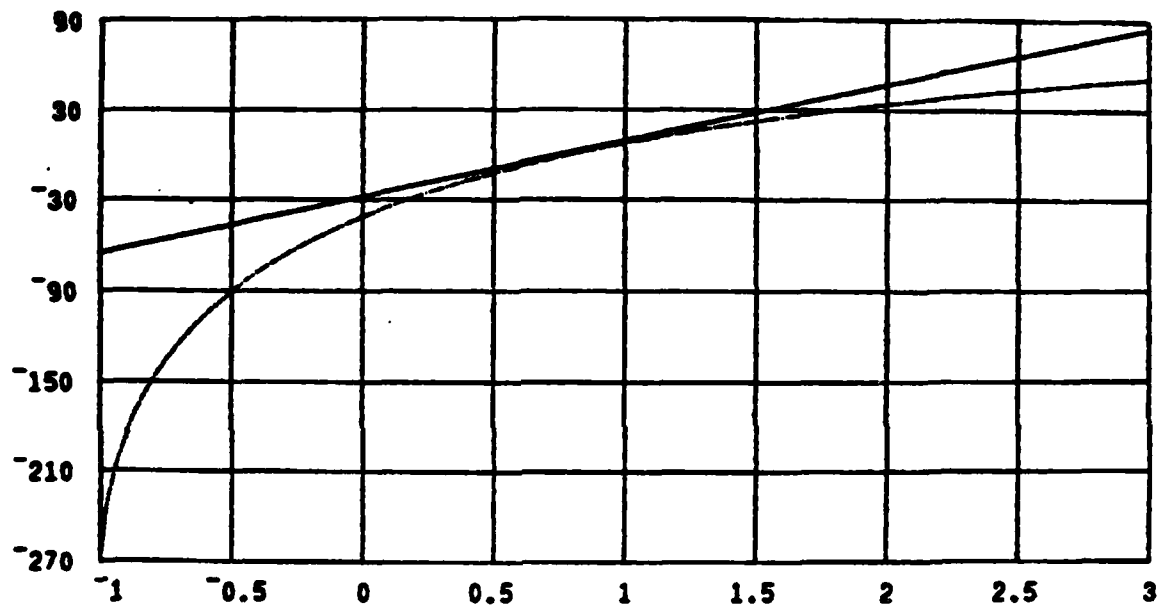
————— Gaussian ($\xi = 0$)

— — — Skewed ($\xi = 0.6$)

Fig. 2.E Neyman-Pearson Optimal Statistics

———— Gaussian

— — — Skewed Noise, $\xi = 0.6$

$$\lim_{\ell \to 0} \alpha = \Phi(-\frac{T_0}{B}) \quad \text{and} \quad \lim_{\ell \to 0} \beta = \Phi(-\frac{T_0-s}{B})$$

which can also be derived easily from Eqs. (31) and (32) by L'Hospital's rule.

Furthermore, it is interesting to note that the false-alarm rate and the power are bounded. By observing that

$$\Phi(x)-\Phi(y) < \Phi(x-y) < \Phi(x+y) < \Phi(x) + \Phi(y), \quad y \geq 0$$

and

$$\Phi(x+y) < \Phi(x-y), \quad y < 0$$

one can show that

$$\Phi(-\frac{T_Q+B}{2A}) \leq \alpha \leq \Phi(-\frac{T_Q-B}{2A})$$

and

$$\Phi(-\frac{Q(T_0-s)+B}{2A}) \leq \beta \leq \Phi(-\frac{Q(T_0-s)-B}{2A})$$

In fact, if $T_0 > s - A = s - \frac{\sigma\xi}{6M}$, the power β is upper bounded by 0.5. Note, however, when the sample size M is large and the skewness measure ξ is small, the case $T_0 > s-A$ is unlikely to happen.

2.6 Conclusions

Two applications of series expansions for detection problems are discussed in this chapter, the Edgeworth series and the Cornish-Fisher inverse expansion. For the former, an approximation to the locally optimum test statistic is evaluated by using the Edgeworth expansion to represent the noise density function. This approximation is not guaranteed to converge to the locally optimum statistic, unless the Tauberian condition is satisfied. Nevertheless, it is applicable to a certain

family of noise distributions since only the first few moments are required to obtain this approximation. Moreover, the scheme thus devised may still yield good detector performance. More sophisticated series expansions may be introduced to achieve better solutions. In the latter part of this chapter, an application of the Cornish-Fisher expansion is demonstrated in the study of the effect of noise skewness on the sample mean detector. The sample mean detector is shown to be sensitive to the noise skewness.

References

[1] D.L. Wallace, "Asymptotic Approximations to Distributions," *Ann. Math. Statist.*, vol. 29, pp. 635-654, 1958.

[2] S.C. Schwartz, "Estimation of Probability Density by An Orthogonal Series," *Ann. Math. Statist.*, vol. 38, 1261-1265, 1967.

[3] L.V. Rozovskii, "On Asymptotic Expansions for The Probability Functions of Sums of Independent Random Vectors," *Theor. Probability and Math. Statist.*, No. 11 pp. 145-154, 1976.

[4] G.L. Anderson and R.J.P. DeFigueiredo, " An Adaptive Orthogonal Series Estimator for Probability Density Functions," *Ann. of Statist.*, vol. 8, pp.347-376, 1980.

[5] H. Cramer, Mathematical Methods of Statistics, Princeton: Princeton University Press, 1946.

[6] M.G Kendall and A. Stuart, The Advanced Theory of Statistics. Vol. I, New York: Hafner, 1963.

[7] R.S. Freedman, "On Gram-Charlier Series," *IEEE Trans. Commun.* vol. COM-29, pp. 122-125, 1981

[8] J. Marcum, "Statistical Theory of Target Detection by Pulsed Radar," *IRE Trans. Inform. Theory* , vol. IT-6, pp. 59-267, 1960.

[9] M. Celebiler and O. Shimbo, "The Probability of Error Due to Intersymbol Interference And Gaussian Noise in Digital Communication Systems," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp.113-119, 1971.

[10] A. Bodharamik *et al.*, "Optimal Detection And Signal Design for Channels with Non-But Near Gaussian Additive Noise," *IEEE Trans. Commun.*, vol. COM-20, pp.1087-1096, 1972.

[11] G. Cariolaro and S. Pupolin, "Considerations on Error Probability in Correlated-Symbol Systems," *IEEE Trans. Commun.*, vol. COM-25, pp. 462-467, 1977.

[12] M. Mansuripur *et al.*, "Fiber Optics Receiver Error Rate Prediction Using The Gram-Charlier Series," *IEEE Trans. Commun.*, vol. COM-28, pp.402-407, 1980.

[13] E.A. Cornish and R.A. Fisher, "Moments And Cumulants in The Specification of Distributions," Paper #31 in R.A. Fisher, Contributions to Mathematical Statistics. New York: John Wiley, 1950.

[14] J.H. Miller and J.B. Thomas, "Detectors for Discrete-Time Signals in Non-Gaussian Noise," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 241-250, 1972.

[15] N.G. De Bruijn, Asymptotic Methods in Analysis, 2nd Edition. New York: Interscience Publishers Inc., 1961.

[16] H.W. Turnbull, Theory of Equations. New York: Interscience Publishers Inc., 1957.

Appendix 2.A

Derivation of Equation (23)

The notations used in the following discussion are defined in Section 4. Suppose that the observation sequence is a sequence of i.i.d. random variables, and that the sample size M is sufficiently large so that the Central Limit Theorem can be applied here. Now, define

$$T(\mathbf{x}) \triangleq \sum_{i=1}^{M} T(x_i)$$

Then the distribution of $T(\mathbf{x})$, by the Central Limit Theorem, is approximately a normal distribution with zero mean and variance $M\sigma_T^2$, where $\sigma_T^2 = Var \ [T(X_i)]$.

The threshold, T_0, and the power, β, of this detection scheme can thus be determined with a prescribed false-alarm rate α as follows :

$$\alpha \approx \frac{1}{\sqrt{2\pi M} \ \sigma_T} \int_{T_0}^{\infty} \exp^{[-x^2/2M\sigma_T^2]} \ dx = 1 - \Phi[\ \frac{T_0}{\sqrt{M} \ \sigma_T}]$$

Hence the threshold for the given false-alarm rate α is equal to

$$T_0 = \sqrt{M} \ \sigma_T \ \Phi^{-1} \ (1-\alpha)$$

Also, the power β is

$$\beta = \Phi[\ \sqrt{M} \ \frac{s}{\sigma_T} - \Phi^{-1}(1-\alpha)]$$

On the other hand, the power β_l of the linear detector using the same threshold and false-alarm rate is given by

$$\beta_l = \Phi[\ \sqrt{N}s - \Phi^{-1}(1-\alpha)]$$

Thus the *ARE* of the nonlinear scheme compared to the linear detector is

$$ARE_{nlod,ld} = \lim_{M,N \to \infty} \frac{M}{N} = \frac{1}{\sigma_T^2}$$

with $s \to 0$, and $\beta = \beta_l$

Appendix 2.B

Discussion of Equation (24)

Let us examine $(2x - h(x))$ first.

$$2x - h(x) = 2x - \frac{\frac{1}{6}\kappa_4 H_3(x) + \frac{1}{120}\kappa_6 H_5(x)}{1 + \frac{1}{24}\kappa_4 H_4(x) + \frac{1}{720}\kappa_6 H_6(x)}$$

$$= \frac{2x[\,1 + \frac{1}{24}\kappa_4 H_4(x) + \frac{1}{720}\kappa_6 H_6(x)\,] - \frac{1}{6}\kappa_4 H_3(x) - \frac{1}{120}\kappa_6 H_5(x)}{1 + \frac{1}{24}\kappa_4 H_4(x) + \frac{1}{720}\kappa_6 H_6(x)}$$

$$= \frac{a_1 x + a_3 x^3 + a_5 x^5 + a_7 x^7}{a_0 + a_2 x^2 + a_4 x^4 + a_6 x^6}$$

where

$$a_0 = 1 + \frac{1}{8}\kappa_4 - \frac{1}{48}\kappa_6$$

$$a_1 = 2 + \frac{3}{4}\kappa_4 - \frac{1}{6}\kappa_6$$

$$a_2 = -\frac{1}{4}\kappa_4 + \frac{1}{16}\kappa_6$$

$$a_3 = -\frac{2}{3}\kappa_4 + \frac{5}{24}\kappa_6$$

$$a_4 = \frac{1}{24}\kappa_4 - \frac{1}{48}\kappa_6$$

$$a_5 = \frac{1}{12}\kappa_4 - \frac{1}{20}\kappa_6$$

$$a_6 = \frac{1}{720}\kappa_6$$

$$a_7 = \frac{1}{360}\kappa_6$$

Consider the following lines located on the (κ_4 κ_6) plane :

$$L_0: \quad 6\kappa_4 - \kappa_6 + 48 = 0$$

$$L_1: \quad 9\kappa_4 - 2\kappa_6 + 24 = 0$$

$$L_2: \quad -4\kappa_4 + \kappa_6 = 0$$

$$L_3: \quad -16\kappa_4 + 5\kappa_6 = 0$$

$L_4:$ $2\kappa_4 - \kappa_6 = 0$

$L_5:$ $5\kappa_4 - 3\kappa_6 = 0$

$L_6:$ $\kappa_6 = 0$

Now, consider, for example, the following noise model :

$$f(x) = (1-\varepsilon)f_1(x) + \varepsilon f_2(x)$$

where

$$f_1(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$$

$$f_2(x) = \frac{\sqrt{2}}{2}e^{-\sqrt{2}|x|}$$

Then clearly,

$$\kappa_4 = 3\varepsilon \qquad ; \qquad \kappa_6 = 30\varepsilon$$

and

$\kappa_4 = \frac{1}{10}\kappa_6$ with $0 \leq \kappa_4 \leq 3$,   (since $0 \leq \varepsilon \leq 1$) .

The lines L_0 to L_6 and their corresponding signed-half spaces are shown on Fig. 2.F. From analytic geometry, one can see that

$\{ a_0, a_1, a_2, a_3, a_6, a_7 \} \geq 0$

and

$\{ a_4, a_5 \} \leq 0$

According to Descartes' law of signs [16], we conclude that both the denominator and numerator will have either no positive real roots (p.r.r.) or two p.r.r.'s. If either the numerator or the denominator has two p.r.r., then $(2x-h(x)) < 0$ holds in some subinterval of $(0, -\infty)$. In this case, $h(x) > 0$, thus we have

$$2xh(x) < h^2(x)$$

On the other hand, if neither the denominator nor the numerator has p.r.r., then $(2x-h(x)) > 0$, for all $x \in [0,\infty]$. Nevertheless, $h(x)$ may be
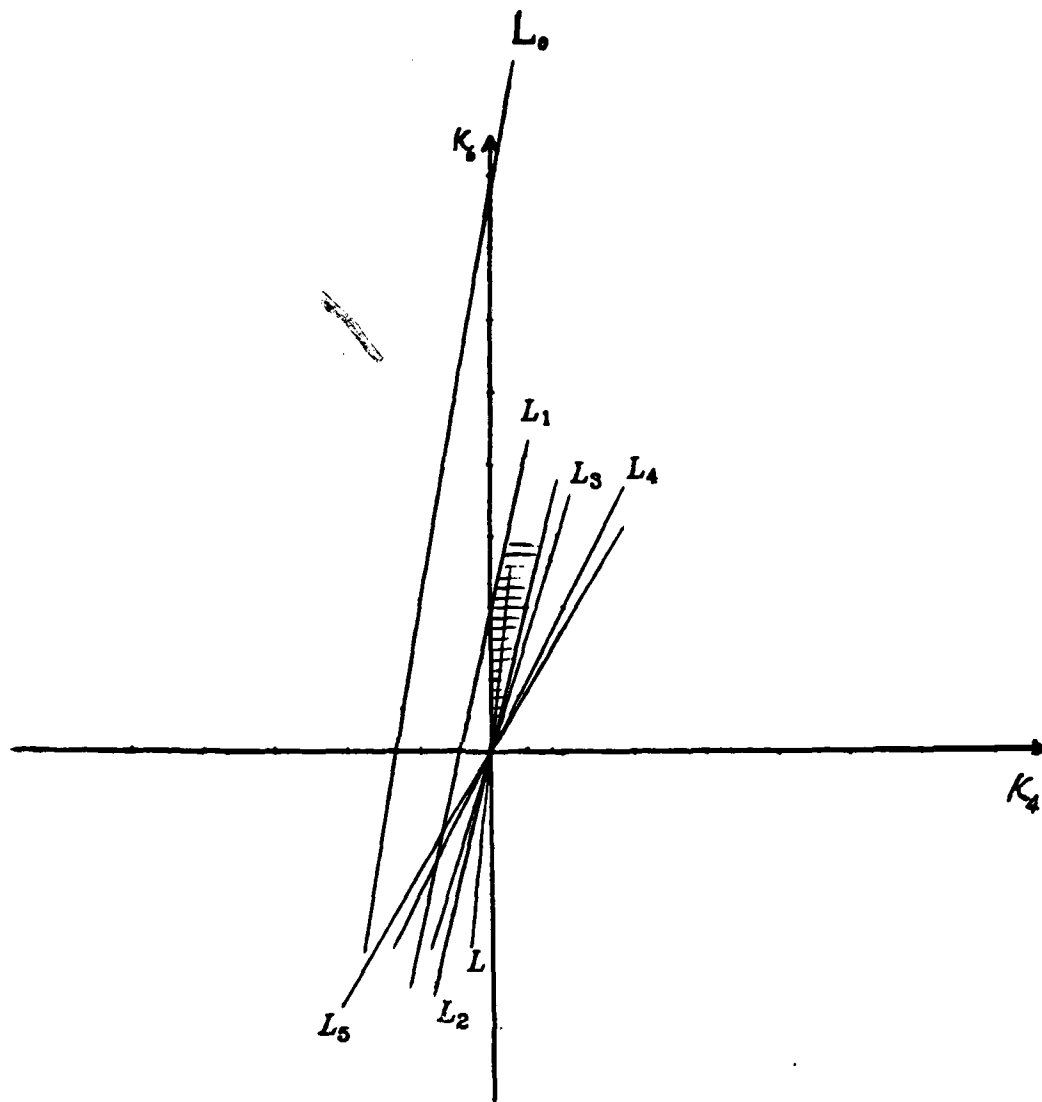
$$L : \kappa_6 = 10 \; \kappa_4$$

Fig. 2.F A Region on Which $(2x - h(x)) > 0$ Holds

(An Example)

negative for some x, so $(2x-h(x))h(x)$ is not always nonnegative even if $2xh(x) < h^2(x)$.

The above discussion actually illustrates one way of investigating the performance of the approximate locally optimum detector using the knowledge on the moments. Further study would be required if a necessary and sufficient condition for $ARE_{alod,ld} > 1$ is desired.

CHAPTER 3

SIGNAL DETECTION IN NEARLY GAUSSIAN SKEWED NOISE

3.1 Introduction

Much, if not most, of the large literature on signal detection has been based on the assumption that the underlying noise is symmetrically distributed with respect to the mean. Typical examples include the sample mean detector and most robust detectors [1] - [3]. This assumption, although generally making problems more tractable and providing simpler solutions, may not be exactly true in practice. In such applications as sonar, asymmetrically distributed noise which might result, for example, from underwater reverberation or sea clutter, has been encountered [4]. Furthermore, data from a number of natural environments, e.g. under-ice ambient noise [5], have actually shown asymmetrical statistical properties. In such situations, optimal detectors could still be found, but there are some additional difficulties in the implementation of such detection schemes. The test statistics are usually complicated or intractable and to implement them precisely requires, of course, a complete knowledge of the underlying noise density functions. Furthermore, the performance of these detection schemes may be sensitive to the inaccuracy of the presumed statistics, and in particular, to the exact amount of skewness.

In a recent paper, Kassam et. al. [6] discussed a robust detection problem using a noise model which allows a symmetric contaminated-nominal central part and an arbitrary tail behavior. Except for this

example, the problem of detection with skewed noise does not seem to have been discussed extensively in the signal detection literature. However, in the context of other statistical applications, there has been some work on investigating the effect of population skewness on hypothesis testings or estimation problems. Johnson [7] discussed the effect of population skewness on the t-variable and proposed a procedure to modify the t-variable so that this effect can be reduced. Carrol [8] examined the effects of asymmetry on estimates of variance of robust estimates in location and regression problems and showed that heavy skewness of errors can seriously bias the commonly used estimates for location and intercept. The purpose of this chapter is to study the effect of noise skewness on the performance of the sample mean detector and the sign detector and, then, to examine a modification of the sample mean detector which is less affected by the noise skewness. Evaluation of the detectors will be based on the false-alarm rate, the power, and the probability of error.

There are many ways to define the skewness measure. Throughout this chapter, this measure is taken to be the third-order central moment divided by the cube of the standard deviation. This definition is one of the most commonly used. By introducing an asymptotic expansion for the test statistics, the detector performance can be related explicitly to this measure. In order to facilitate analysis, the test statistics will be normalized by the sample size M. This normalization procedure will not affect the performance of the detectors so long as the thresholds are adjusted correspondingly. Then the relation between the performance

and the noise skewness is obtained. This relation will give a clear qualitative insight into detector performance, although, for any finite M, it will be an approximation. It will turn out that the sample mean detector is more sensitive to noise skewness than the sign detector, when the underlying noise is only sightly non-Gaussian due to skewness. This result is reasonable since the sign detector only assumes a zero median, a milder assumption than symmetry, for the noise distribution. The second part of this chapter then proposes a modified scheme for the sample mean detector, which, by the introduction of a correction term, reduces the effect of noise skewness.

Curves which describe the performance change with respect to noise skewness measure for both the sample mean detector and the sign detector are given in Figs. 3.A - 3.C. Specifically, Fig. 3.A shows rates of change in the false-alarm rate for both detectors, Fig. 3.B depicts those in the power, and Fig. 3.C is for the probability of error. Note that, in these figures, the thresholds are kept constant as the skewness measure varies. Fig. 3.D illustrates the structure of the modified sample mean detector. In the last part of this chapter, some numerical examples based on Monte-Carlo simulations and on some data from under-ice ambient noise are given.

3.2 Preliminary Results

One of the major difficulties in examining conventional detectors with asymmetrically distributed noise has been that the skewness measure does not usually appear explicitly in the test statistics and thus perfor-

mance of the detector can hardly be expressed in terms of this measure. Although several different skewness measures are available in the literature [9],[10], this difficulty remains unchanged. The reason appears to be that, in general, skewness measures are evaluated as a function of the third-order moment, a function of the mean and the mode or the median, or even as a function of the kurtosis and the third-order moment. On the other hand, the test statistics usually evolve from the probability density functions (p.d.f.) or from other functions of the distribution which may not be related explicitly to those skewness measures mentioned above. Hence, even though it is intuitively clear that the statistical properties of the test statistics are affected by the noise skewness, the explicit functional relations are hard to acquire. One reasonable way to get around this problem is to obtain an asymptotic expansion for the test statistic in terms of the moments which are more clearly related to the usual skewness measures.

The main purpose of this section is to develop some preliminary results on which the procedure of expanding the test statistics will be based. There are several general ways to obtain an asymptotic expansion for a random variable. The one to which attention here will be restricted is the use of the Cornish-Fisher expansion [11]. A general form of such an expansion for a random variable X is given by

$$CF(X) = \mu_0 + \sigma Z + \frac{\mu_3}{\sigma^2}(Z^2 - 1) + \cdots \tag{1}$$

where $\mu_0 = E\{X\}$, $\sigma^2 = E\{(X-\mu_0)^2\}$, $\mu_3 = E\{(X-\mu_0)^3\}$, and Z is a standard normal random variable. In this chapter, our interest is to employ this sort of expansion for the sample mean of a sequence of independent and

identically distributed (i.i.d.) random variables so that some asymptotic properties can be achieved. To begin with, the asymptotic expansion for the sample mean should be defined. Let $\{X_i\}$ be a sequence of i.i.d. random variables with sample mean $S_M = \frac{1}{M} \sum_{i=1}^{M} X_i$, where M denotes the sample size. We will call a series of random variables, $a_0 + \sum_{i=1}^{r} \frac{a_i}{M^{i/2}} Z_i$, an "asymptotic expansion valid to r terms" for S_M if

$$|S_M - a_0 - \sum_{i=1}^{r} \frac{a_i}{M^{i/2}} Z_i| = o(M^{-r/2}) \quad w.p. 1. \tag{2}$$

This definition evolves from Wallace [12] who defined an asymptotic expansion valid to r terms for a distribution function while requiring the remainder to be $O(M^{-r/2})$. However, it will be seen later that requiring the remainder to be $o(M^{-r/2})$ as in Eq.(2) is more convenient and furthermore, is in accordance with Erdelyi [13]. It is clear that, when $r = 1$, if a_0 is taken to be the mean of X_i and Z_1 is the standard normal random variable, then Eq.(2) is compatible with the ordinary Central Limit Theorem. Furthermore, one can argue that the Central Limit Theorem suggests that, under some general conditions, the distribution of S_M tends to be symmetric. However, it will be shown here that, by introducing the series expansion for the test statistics, one can obtain a solution with faster convergence rate. Now, let σ^2 and μ_3 be the variance and the third-order central moment of X_i respectively, and let the skewness measure ξ be defined by

$$\xi = \frac{\mu_3}{\sigma^3} \tag{3}$$

Then, it is easy to see that the skewness measure ξ_M of the distribution of

S_M is given by

$$\xi_M = \frac{1}{\sqrt{M}} \cdot \frac{\mu_3}{\sigma^3} \to 0 \ as \ M \to \infty$$

These arguments lead to the following lemma:

Lemma: Let $\{X_i\}$ be a sequence of i.i.d. random variables with common distribution F. Assume that all moments of X_i exist and that the skewness measure of F is defined by Eq.(3). Then the skewness measure ξ_M of the distribution of S_M satisfies the following relation

$$\xi_M = O(M^{-1/2})$$

However, the interest here is to consider an asymptotic expansion valid to higher order terms such that the expansion will contain the skewness measure. In fact, these higher order terms are important since taking one or two more terms usually improves the approximation significantly and, typically, may correct the skewness. Further investigation reveals that Eq.(2) determines an equivalence relation between sequences of i.i.d. random variables. We can say that two sequences of i.i.d. random variables are asymptotically equivalent if their asymptotic expansions valid to r terms, for the sample mean, differ by $o(M^{-r/2})$ for each r. Thus an equivalence class based on this relation may be defined. And therefore, a valid asymptotic expansion defined in Eq.(2) may represent a class of sequences of i.i.d. random variables.

We now proceed to present a Cornish-Fisher type expansion for the sample mean of a sequence of i.i.d. random variables $\{X_i\}$. The following proposition is a straightforward result :

Proposition 1: Let $\{X_i\}$ be a sequence of i.i.d. random variables and S_M be its size-M sample mean. Assume that all moments of X_i exist and that $|\frac{\mu_3}{\sigma^3}| \ll 1$; then an expansion for S_M given by

$$VAE(S_M) = \mu_0 + \frac{\sigma}{\sqrt{M}}Z + \frac{\mu_3}{6\sigma^2 M}(Z^2-1) \qquad (4)$$

is an asymptotic expansion valid to two terms for S_M, where $\mu_0 = E\{X_i\}$, $\sigma^2 = E\{(X_i - \mu_0)^2\}$, $\mu_3 = E\{(X_i - \mu_0)^3\}$, and Z is a standard normal random variable.

If the skewness measure is defined by Eq. (3), then Eq. (4) obviously contains this measure. Now, by the use of some weak convergence theorems on pages 287 and 288 in [14] and the observation that $VAE(S_M)$ is a continuous function of Z, *Proposition 2* follows immediately from *Proposition 1*.

Proposition 2: Let $Y_M = \mu_0 + \frac{\sigma}{\sqrt{M}}Z + \frac{\mu_3}{6\sigma^2 M}(Z^2-1)$. Let F_{Y_M} and F_{S_M} be the distribution functions of Y_M and S_M respectively. Then, under the assumptions in *Proposition 1*, Y_M converges in law to S_M ; namely,

$$\lim_{M \to \infty} F_{Y_M} = F_{S_M}$$

at each continuity point of F_{S_M}.

Proposition 2 enables us to evaluate the asymptotic statistics of S_M via those of Y_M. As a matter of fact, the p.d.f. of Y_M exists and can be evaluated as follows:

Substituting Eq. (3) into Eq. (4) yields

$$Y_M = VAE(S_M) = \mu_0 + \frac{\sigma}{\sqrt{M}}Z + \frac{\xi\sigma}{6M}(Z^2-1) \qquad (5)$$

Then, it can be shown that

(i)If $\xi \geq 0$, the p.d.f of Y_M is given by

$$\frac{dF_{Y_M}(y)}{dy} \begin{cases} = \sqrt{\frac{M}{2\pi\sigma^2}} Q^{-1}(y) \{\Phi(\frac{-1}{\sigma_l}) \exp[-\frac{1}{2}(\frac{Q(y)+1}{\sigma_l})^2] + \\ \quad \Phi(\frac{1}{\sigma_l}) \exp[-\frac{1}{2}(\frac{Q(y)-1}{\sigma_l})^2]\} \quad \text{if } y \geq \mu_0 - \frac{3\sigma}{2\xi} - \frac{\sigma\xi}{6M} \quad (6) \\ = 0 \qquad\qquad\qquad\qquad \text{otherwise} \end{cases}$$

where

$$Q(y) = \sqrt{1 + \sigma_l^2 + 2\xi(y-\mu_0)/3\sigma} \quad \text{and} \quad \sigma_l = \xi/3\sqrt{M}$$

(ii) If $\xi < 0$, the p.d.f. of Y_M is given by

$$\frac{dF_{Y_M}(y)}{dy} \begin{cases} = \sqrt{\frac{M}{2\pi\sigma^2}} Q^{-1}(y) \{\Phi(\frac{-1}{\sigma_l}) \exp[-\frac{1}{2}(\frac{Q(y)-1}{\sigma_l})^2] + \\ \quad \Phi(\frac{1}{\sigma_l}) \exp[-\frac{1}{2}(\frac{Q(y)+1}{\sigma_l})^2]\} \quad \text{if } y \leq \mu_0 - \frac{3\sigma}{2\xi} - \frac{\sigma\xi}{6M} \quad (7) \\ = 0 \qquad\qquad\qquad\qquad \text{otherwise} \end{cases}$$

The preceding discussions have been focused on continuous random variables. It should be noted that these expansions may not be valid for some discrete random variables. However, in the context of this chapter, only binomially distributed random variables will be discussed and these will be seen to have valid asymptotic expansions. Thus the discussion for the discrete case will be omitted here.

3.3 The Sample Mean Detector

The sample mean detector is known to be the Neyman-Pearson optimal detector when the underlying noise density is Gaussian. This section considers sample mean detector performance deterioration under model deviation due to skewness. The asymptotic expansion introduced in the previous section will be employed here for the test statistic, and

relations between the performance and the noise skewness measure will then be obtained. It will be shown that the performance of the sample mean detector does deteriorate, as expected.

The following binary hypothesis testing problem is considered:

$$H_0 : X_i = N_i$$
$$K_0 : X_i = N_i + s \qquad i=1,2,...,M$$

It is assumed here that $\{X_i\}$ is a sequence of i.i.d. random variables. Under H_0, the p.d.f. of X_i is $f(x_i)$, which is nearly Gaussian, and has zero mean, known variance σ^2 and known third-order moment μ_3, where μ_3 is small. Now, the test statistic of the sample mean detector is given by

$$T(X) = \frac{1}{M}\sum_{i=1}^{M} X_i$$

As discussed in Section 3.2, there exists an asymptotic expansion valid to two terms for $T(X)$. For notational convenience, the subscript M will be omitted in the sequel and the asymptotic expansion for $T(X)$ will be denoted by Y. Thus, in accordance with Eq. (5),

$$\text{under } H_0 : Y = Y_H = \frac{\sigma}{\sqrt{M}}Z + \frac{\sigma\xi}{6M}(Z^2-1) \qquad (8)$$

$$\text{under } K_0 : Y = Y_K = s + \frac{\sigma}{\sqrt{M}}Z + \frac{\sigma\xi}{6M}(Z^2-1) \qquad (9)$$

Let $F_H(y)$ and $F_K(y)$ denote the distribution functions of Y_H and Y_K, respectively. Now, define α_M and β_M by

$$\alpha_M = \int_{T_0}^{\infty} dF_H(y) \qquad (10)$$

and

$$\beta_M = \int_{T_0}^{\infty} dF_K(y) \qquad (11)$$

where T_0 is the given threshold. Then, by *Propositions* 1 and 2, it can be concluded that α_M and β_M converge respectively to the false-alarm rate

α_l and the power β_l of the sample mean detector. Now,

(i) if $\xi \geq 0$, substituting Eq. (6) with $\mu_0 = 0$ and $\mu_0 = s$ into Eqs. (10) and (11) respectively yields

$$\alpha_M = \Phi(-\frac{1}{\sigma_l})\Phi(-\frac{Q_0(T_0)+1}{\sigma_l}) + \Phi(\frac{1}{\sigma_l})\Phi(-\frac{Q_0(T_0)-1}{\sigma_l}) \qquad (12)$$

and

$$\beta_M = \Phi(-\frac{1}{\sigma_l})\Phi(-\frac{Q_s(T_0)+1}{\sigma_l}) + \Phi(\frac{1}{\sigma_l})\Phi(-\frac{Q_s(T_0)-1}{\sigma_l}) \qquad (13)$$

where

$$Q_0(T_0) = \sqrt{1+\sigma_l{}^2+2\xi T_0/3\sigma} \quad \text{and} \quad Q_s(T_0) = \sqrt{1+\sigma_l{}^2+2\xi(T_0-s)/3\sigma}$$

When M is sufficiently large and ξ is small, α_M and β_M can be approximated by

$$\alpha_M \approx \Phi(-\frac{Q_0(T_0)-1}{\sigma_l}) \qquad (14)$$

and

$$\beta_M \approx \Phi(-\frac{Q_s(T_0)-1}{\sigma_l}) \qquad (15)$$

(ii) if $\xi < 0$, substituting Eq. (7) with $\mu_0 = 0$ and $\mu_0 = s$ into Eqs. (10) and (11) respectively yields

$$\alpha_M = \Phi(\frac{-1}{\sigma_l})[\Phi(-\frac{Q_0(T_0)-1}{\sigma_l}) - \Phi(\frac{1}{\sigma_l})] +$$

$$\Phi(\frac{1}{\sigma_l})[\Phi(-\frac{Q_0(T_0)+1}{\sigma_l}) - \Phi(\frac{-1}{\sigma_l})] \qquad (12a)$$

and

$$\beta_M = \Phi(\frac{-1}{\sigma_l})[\Phi(-\frac{Q_s(T_0)-1}{\sigma_l}) - \Phi(\frac{1}{\sigma_l})] +$$

$$\Phi(\frac{1}{\sigma_l})[\Phi(-\frac{Q_s(T_0)+1}{\sigma_l}) - \Phi(\frac{-1}{\sigma_l})] \qquad (13a)$$

Similarly, when M is sufficiently large and ξ is small, α_M and β_M can be approximated by Eqs. (14) and (15) respectively. Since α_M and β_M converge to α_l and β_l respectively, the qualitative behavior of α_l and β_l

can be studied approximately by α_M and β_M respectively, provided that M is sufficiently large. Now the probability of error P_{el} is given by

$$P_{el} \approx \Phi(\frac{Q_s(T_0)-1}{\sigma_l}) + \Phi(-\frac{Q_0(T_0)-1}{\sigma_l}) \qquad (16)$$

It can be seen from Eqs. (14)-(16) that the dependence of the detector performance on the noise skewness measure comes from the variables $Q_0(T_0)$, $Q_s(T_0)$, and σ_l. Figs. 3.A - 3.C provide clearer quantitative insights as to sensitivity of the sample mean detector to the underlying noise skewness. One can see that both the false-alarm rate and the probability of error increase as the skewness measure increases, while the power stays nearly the same, an indication of the performance deterioration of the sample mean detector when the noise skewness comes into play.

Before concluding this section, it may be worth mentioning that, throughout the above discussion , we have assumed implicitly that the threshold T_0 lies inside the region where the corresponding p.d.f. for Y_M is nonzero. The case when T_0 lies outside that region will result in a singular detection problem which is of little interest here.

3.4 The Sign Detector

It is well known that the sign detector is the Neyman-Pearson optimal detector for the nonparametric test (H_1, K_1)

$$H_1 : p = 1/2$$
$$K_1 : p > 1/2$$

where $p = Prob. \{ X_i \geq 0 \}$. Also, it is the locally optimal detector when the underlying noise has a Laplace distribution [15],[16]. The
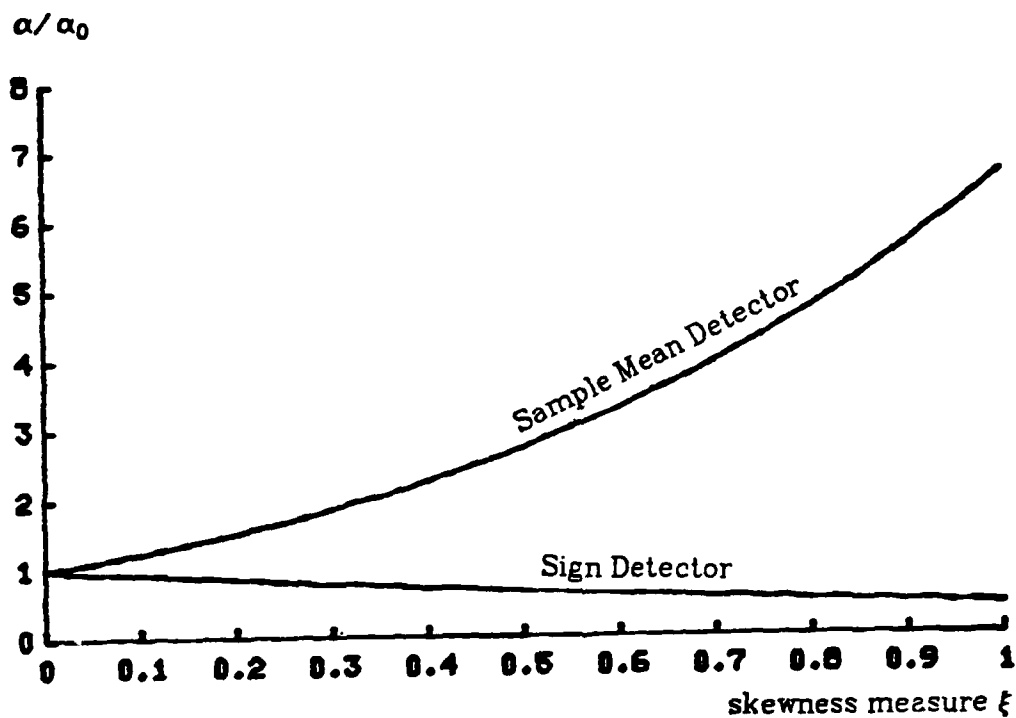
Fig. 3.A Rate of Change in False-alarm Rate

α_0 is the false-alarm rate when the underlying noise is Gaussian

$$\alpha_0 = 1.0 \times 10^{-10}$$

The threshold is set by α_0 and is kept fixed when the

noise skewness measure ξ varies.

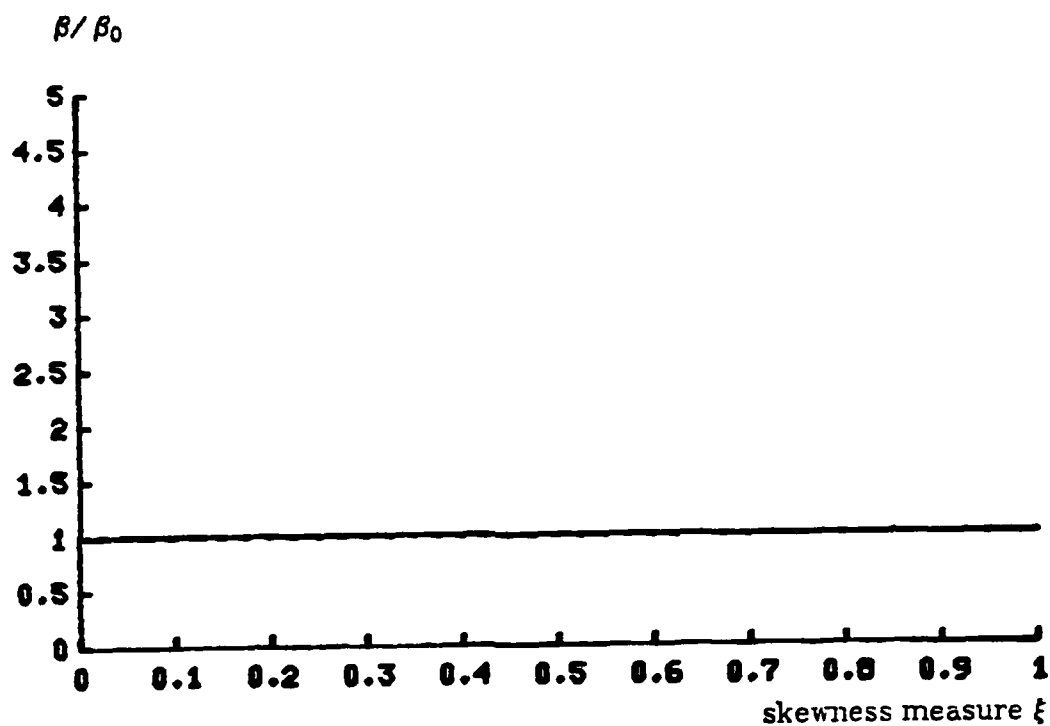Sample size = 400

β / β_0

Fig. 3.B Rate of Change in Power

β_0 is the power when the noise is Gaussian

$$\beta_0 \begin{cases} \approx 0.99999 & \text{for the linear detector} \\ \approx 0.99998 & \text{for the sign detector} \end{cases}$$

Signal Strength = 2 × (threshold of the linear detector)
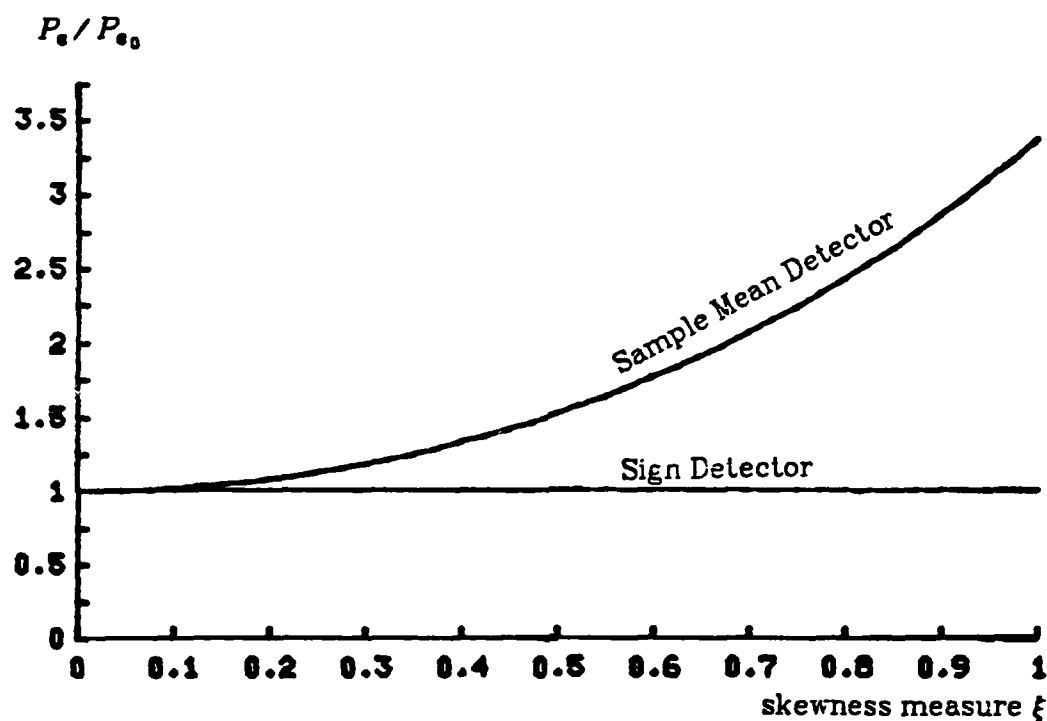
Sample size = 400

Fig. 3.C Rate of Change in Probability of Error

P_{e_0} is the probability of error when the noise is Gaussian

$$P_{e_0} \begin{cases} = 2.0 \times 10^{-10} & \text{for the linear detector} \\ \approx 2.0 \times 10^{-5} & \text{for the sign detector} \end{cases}$$

Sample size = 400

performance of the sign detector when the underlying noise is skewed will now be studied. It is easy to see that, when the underlying noise is skewed, the binary test (H_1, K_1) really becomes

$$H_2 : p = 1/2 + \varepsilon(\xi)$$
$$K_2 : p > 1/2 + \varepsilon(\xi)$$

where $\varepsilon(\xi)$ denotes the deviation depending on the skewness measure ξ. In order to be compatible with the previous discussion, the following normalized test statistic is considered:

$$T(X) = \frac{1}{M} \sum_{i=1}^{M} sgn(X_i) \tag{17}$$

where

$$sgn(X_i) = \begin{cases} 1 & \text{if } X_i \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Again, an asymptotic expansion for $T(X)$ will be needed in order to facilitate analysis. In a survey paper, Bickel [17] discussed the Edgeworth expansion in nonparametric statistics with emphasis on higher-order approximations to the distribution of those statistics. However, the test statistic of the sign detector does not seem to have been discussed. In this section, the previously considered Cornish-Fisher type asymptotic expansion will be utilized again for $T(X)$. The first three central moments of $T(X)$ under H_2 and K_2 are given in the following:

$$\text{under } H_2 : \quad \mu_0 = 1/2 + \varepsilon(\xi)$$
$$\sigma^2 = 1/4 - \varepsilon^2(\xi) \tag{18}$$
$$\text{and} \quad \mu_3 = -2\varepsilon(\xi)\sigma^2$$

$$\text{under } K_2 : \quad \mu_0 = p = 1/2 + \varepsilon_s(\xi)$$
$$\sigma^2 = p - p^2 = 1/4 - \varepsilon_s{}^2(\xi) \tag{19}$$
$$\text{and} \quad \mu_3 = p - 3p^2 + 2p^3 = -2\varepsilon_s'(\xi)\sigma^2$$

where, for convenience, p is taken to be $1/2 + \varepsilon_s(\xi)$ under K_2, and $\varepsilon_s(\xi)$ is

a variable depending on both the signal strength and the skewness measure. Now, following the discussion in the previous sections, there exist asymptotic expansions valid to two terms for $T(X)$ under both hypotheses, namely,

$$\text{under } H_2 : Y_{H_2} = 1/2 + \varepsilon(\xi) + \sqrt{\frac{1/4 - \varepsilon^2(\xi)}{M}} Z - \frac{\varepsilon(\xi)}{3M}(Z^2 - 1)$$

$$\text{under } K_2 : Y_{K_2} = 1/2 + \varepsilon_s(\xi) + \sqrt{\frac{1/4 - \varepsilon_s^2(\xi)}{M}} Z - \frac{\varepsilon_s(\xi)}{3M}(Z^2 - 1)$$

Then, by going through the same procedures as of Section 3.3, we can obtain the false-alarm rate α_s and the power β_s for the sign detector.

$$\alpha_s \approx \Phi(\frac{1 - R(T_0)}{\sigma_s}) \tag{20}$$

and

$$\beta_s \approx \Phi(\frac{1 - R'(T_0)}{\sigma'_s}) \tag{21}$$

where

$$\sigma_s = \frac{-2\varepsilon(\xi)}{3\sqrt{M(1/4 - \varepsilon^2(\xi))}} \quad ; \quad \sigma'_s = \frac{-2\varepsilon_s(\xi)}{3\sqrt{M(1/4 - \varepsilon_s^2(\xi))}}$$

and

$$R(T_0) = \sqrt{1 + \sigma_s^2 + \frac{4\varepsilon(\xi)(T_0 + \varepsilon(\xi) - 1/2)}{3(1/4 - \varepsilon^2(\xi))}}$$

$$R'(T_0) = \sqrt{1 + \sigma_s'^2 + \frac{4\varepsilon_s(\xi)(T_0 + \varepsilon_s(\xi) - 1/2)}{3(1/4 - \varepsilon^2(\xi))}}$$

Thus the probability of error P_{es} is

$$P_{es} \approx \Phi(\frac{R'(T_0) - 1}{\sigma'_s}) - \Phi(\frac{1 - R(T_0)}{\sigma_s}) \tag{22}$$

Eqs. (20)-(22) reveal how the performance of the sign detector changes as the underlying noise skewness changes. The dependence on the skewness is embedded in the variables σ_s, σ'_s, $R(T_0)$ and $R'(T_0)$. Figs. 3.A - 3.C also depict the quantitative relation between the performance of the sign detector and the noise skewness. The noise model used here is the same

as that for the sample mean detector. The noise p.d.f. is given in Eq. (6) with $\mu_0 = 0$. For this particular model, which is nearly Gaussian, the sign detector appears to be less sensitive to the noise skewness. This result is reasonable since the sign detector only assumes a zero median of the noise distribution, which is a milder assumption than zero skewness. On the other hand, the sample mean detector assumes the Neyman-Pearson optimality for a Gaussian noise model, which has zero skewness measure.

3.5 A Modified Sample Mean Detector

In Section 3.3, the sample mean detector was examined with skewed noise. The noise model used there is a Gaussian model perturbed with a small amount of skewness. It has been shown that the performance of *the sample mean detector deteriorates* when the underlying noise becomes skewed, although the significance of this performance deterioration depends on the particular application. It is then interesting to consider reducing this effect by modifications of the detector structure. In this section, a modified scheme is proposed based on the asymptotic expansion for the sample mean of an i.i.d. random sequence discussed in Section 3.2. A mean-squared term is added to the sample mean so that, in the asymptotic expansion, the terms involving the lowest power of the skewness measure can be eliminated. Now, with the assumption that the skewness measure is small and that the sample size M is sufficiently large, one can expect that the effect of the noise skewness is essentially eliminated. Again, the following hypothesis testing problem is considered:

$$H_0 : X_i = N_i$$
$$K_0 : X_i = N_i + s \quad i = 1, 2, ..., M$$

As in Section 3.3, it is assumed here that, under H_0, $E\{X_i\}=0$, $E\{X_i{}^2\}=\sigma^2$, and $E\{X_i{}^3\}=\mu_3$, where μ_3 is small. Then, consider a test with the following test statistic:

$$T_m(X) = \bar{X} + \rho(\bar{X}^2 - \frac{\sigma^2}{M}) \tag{23}$$

where

$$\bar{X} = \frac{1}{M}\sum_{i=1}^{M} X_i$$

The parameter ρ will be determined as follows:

By *Proposition* 1, an asymptotic expansion valid to two terms for $\bar{X}$ exists, namely

$$\bar{X} = \frac{\sigma}{\sqrt{M}}Z + \frac{\mu_3}{6\sigma^2 M}(Z^2-1) + o(M^{-1}), \quad \text{under } H_0 \tag{24}$$

Substituting Eq. (24) into Eq. (23) yields

$$T_m(X) = \frac{\sigma}{\sqrt{M}}Z + \frac{\mu_3}{6\sigma^2 M}(Z^2-1) + \rho\frac{\sigma^2}{M}(Z^2-1) + o(M^{-1})$$

Thus, if $\rho = -\mu_3/6\sigma^4$, the term involving μ_3 will be eliminated and other terms involving higher powers of the skewness measure will be contained in $o(M^{-1})$, i.e.

$$T_m(X) = \frac{\sigma}{\sqrt{M}}Z + o(M^{-1})$$

Therefore, when the skewness measure, $\xi = \mu_3/\sigma^3$, is small and the sample size M is sufficiently large, $T_m(X)$ can be well approximated by a normal random variable which is independent of the skewness measure of the noise. Hence the false-alarm rate α_m and the power β_m are given by

$$\alpha_m \approx 1 - \Phi(\sqrt{M}T_0/\sigma)$$

and

$$\beta_m \approx 1 - \Phi(\sqrt{M}\,(T_0 - s)/\sigma)$$

where T_0 is the prescribed threshold. Thus it can be seen here that the performance of the modified sample mean detector is asymptotically independent of the noise skewness. From Eq. (23) and the above discussion, the modified sample mean detector should be implemented by the following test statistic

$$T_m(X) = \bar{X} - \frac{\xi}{6\sigma}(\bar{X^2}) \tag{25}$$

The structure of this modified scheme is shown in Fig. 3.D. In practice, the skewness measure ξ may not be known exactly. Then a learning procedure or adaptive scheme will be needed to obtain an estimate.

3.6 Simulation Examples

To investigate the modified sample mean detector proposed in the last section in more detail, some numerical examples based on Monte-Carlo simulations are given here. For simplicity, the noise model which has a probability density function given by Eq. (6) is used. This is a model resulting from a Gaussian distribution perturbed by a non-zero skewness measure. According to the results discussed in the previous sections, the modified scheme is expected to be less sensitive to noise skewness, compared to the sample mean detector. As a matter of fact, it should be indifferent asymptotically to noise skewness. The following examples verified this conclusion. Fig. 3.E shows a result of one million runs with a sample size of 100 and the α_0, the initial false-alarm rate when the skewness measure ξ is zero, set equal to 10^{-3}. Curves are plotted for the rate
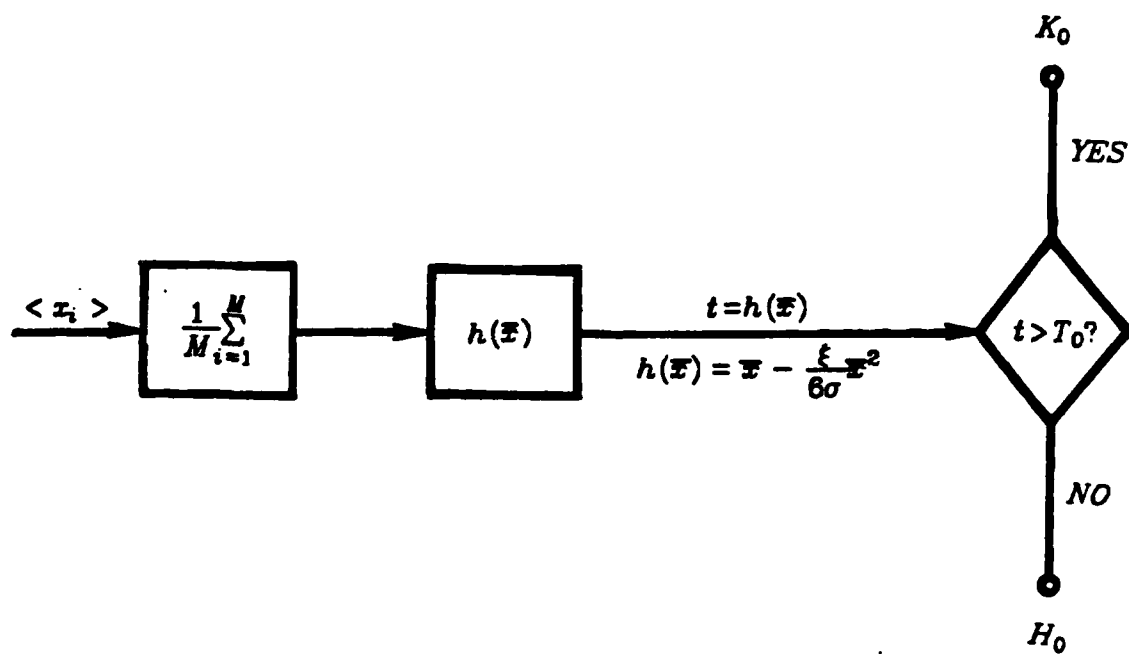
Fig. 3.D A Modified Sample Mean Detector

of change in false-alarm rate as the skewness measure varies. It can be seen that the false-alarm rate of the modified scheme stays nearly unchanged, compared to that of the sample mean detector, as ξ changes. In Fig. 3.F, a similar set of curves is shown, except that $\alpha_0 = 10^{-4}$. It should be noted here that the smoothness of the curves depends on the number of runs. This is one of the typical properties of the Monte-Carlo simulations. However, we do see a positive verification of the analytical results obtained previously.

3.7 A Case Study Using Under-Ice Ambient Noise

The detection scheme discussed in Section 3.5 will be further examined here with some real data taken from under ice in the Arctic. The data were preliminarily analyzed and presented in [5]. The mechanism for the noise under stationary shore ice is possibly due to tensile stresses caused by rapid reduction in air temperatures. Noise from the pack ice, on the other hand, is due to the friction between interacting and colliding ice floes in addition to tensile stresses. For completeness of discussion, the first three statistical moments of the data are shown here in Figs. 3.G - 3.I [5] in the time domain. These moments were estimated in 1024 sample blocks, which were identified as a record on the horizontal axes in these figures. The samples were taken at a sampling rate of $10kHz$, thus each record represents approximately 0.1 second of time. It is seen that the data has a non-zero mean which is due to the carrier frequency of the tape recorder being slightly misaligned. Fig. 3.H shows a large variability of the variance over time. This illustrates the non-stationarity of the
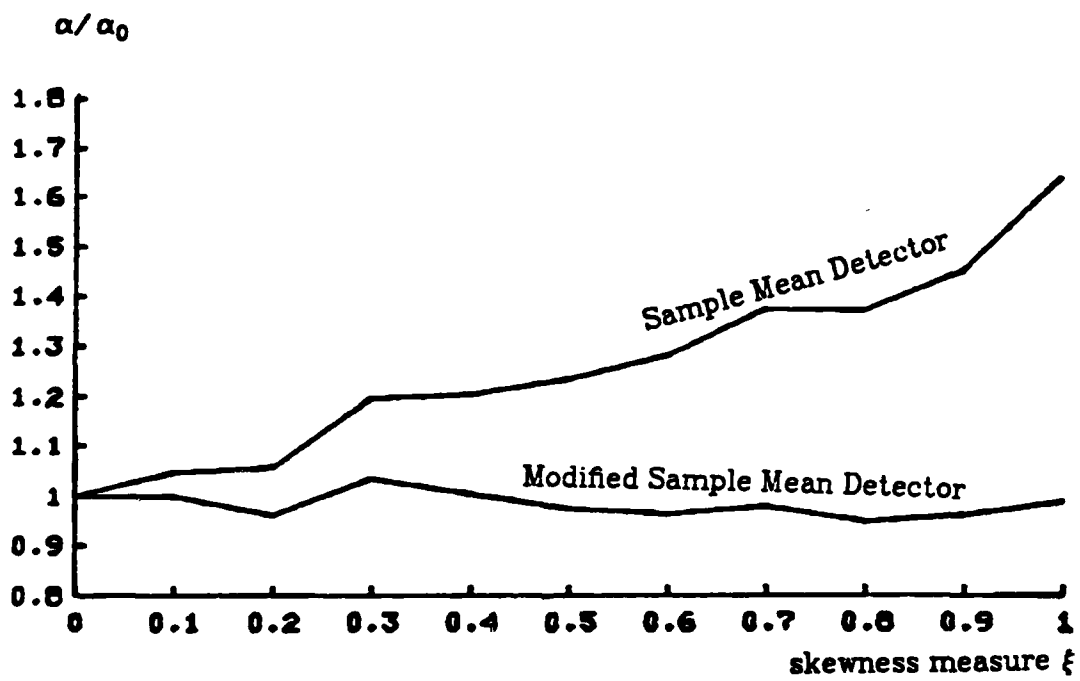
A Result of Monte Carlo Simulation

Fig. 3.E Number of Runs = 10^6

$$\alpha_0 = 10^{-3}$$

Sample Size M = 100
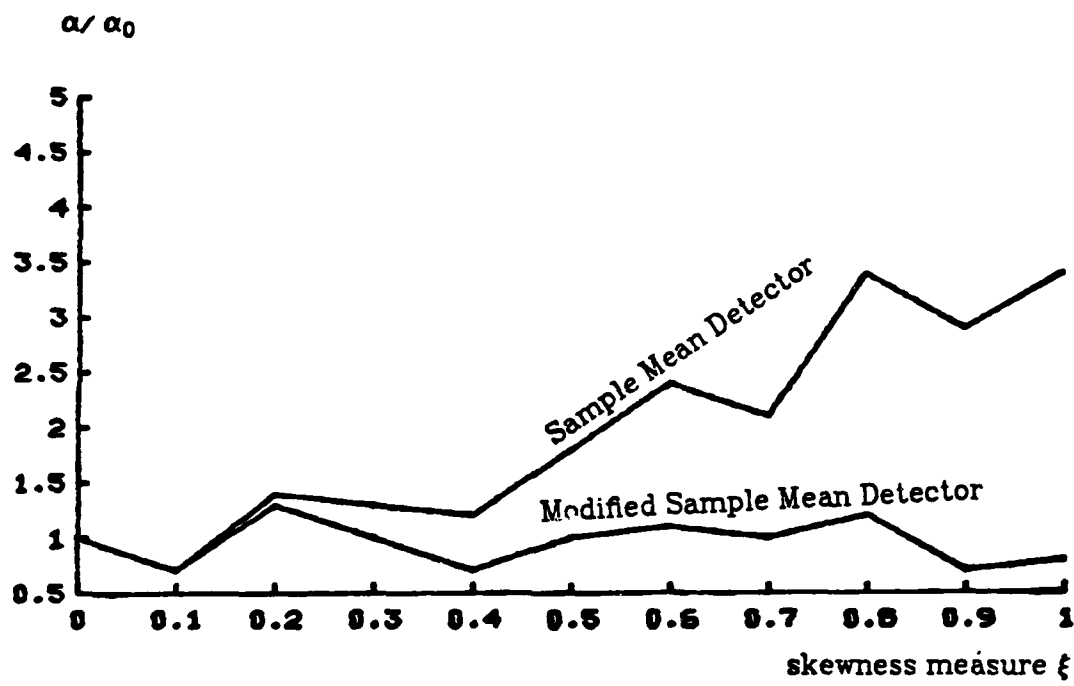
A Result of Monte Carlo Simulation

Fig. 3.F Number of Runs = 10^6

$$\alpha_0 = 10^{-4}$$
Sample Size M = 100

noise. Fig. 3.I depicts the skewness measure which is defined in Eq. (3). It can be seen that, for some records, the noise exhibits a pronounced asymmetrical property.

The purpose of this study is to investigate two detectors (the sample mean detector and the modified one) with data from some natural environment. There are some disadvantages associated with the use of these data, however. As has been seen from Fig. 3.H, the data represent a nonstationary process and, furthermore, they may well represent a dependent process. These two characteristics seriously violate the assumptions discussed in this chapter. Hence the results thus obtained should not be used to justify the analytical results of the previous sections. In an attempt to eliminate the dependency structure in the noise, the data sequence was re-arranged such that every other eight consecutive samples is used. The estimates of the statistical moments are then evaluated based on the re-arranged data sequence. Thus, each record now only consists of 128 sample blocks. Figs. 3.J and 3.K plot the false-alarm rate and the probability of error versus the threshold for the two detectors using this re-arranged data sequence and for the sample mean detector in Gaussian noise. It is shown here, based on these data, that the discrimination between these two detectors does not appear to be significant. It should be noted here that the estimates of the moments obtained from this data sequence does not seem to be different from those shown in Figs. 3.G - 3.I.
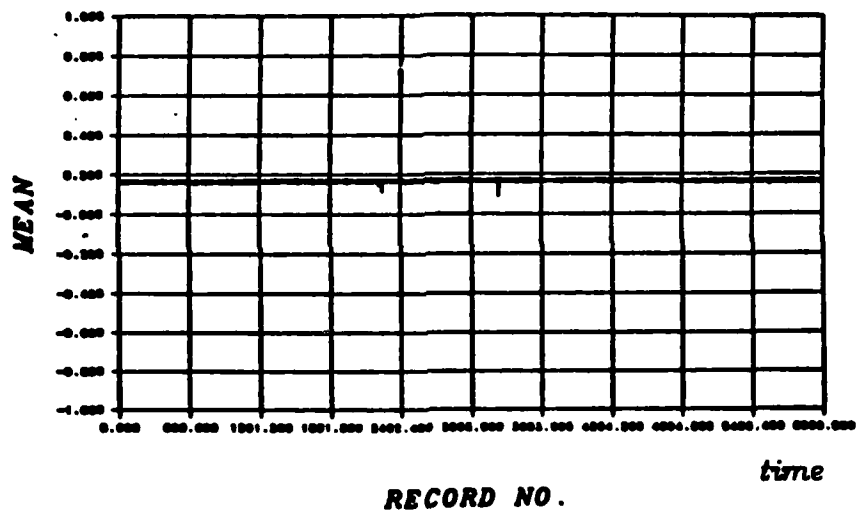
Fig. 3.G Statistical Moment of Under-Ice Noise (Mean)
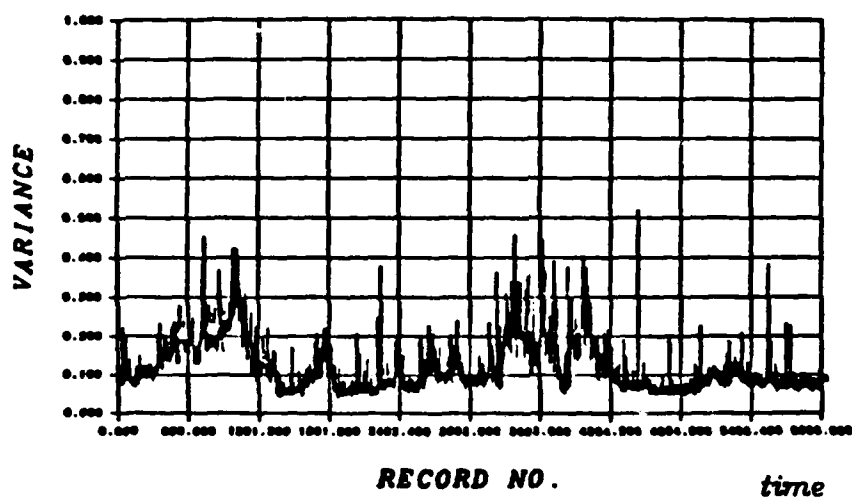
Fig. 3.H Statistical Moment of Under-Ice Noise (Variance)
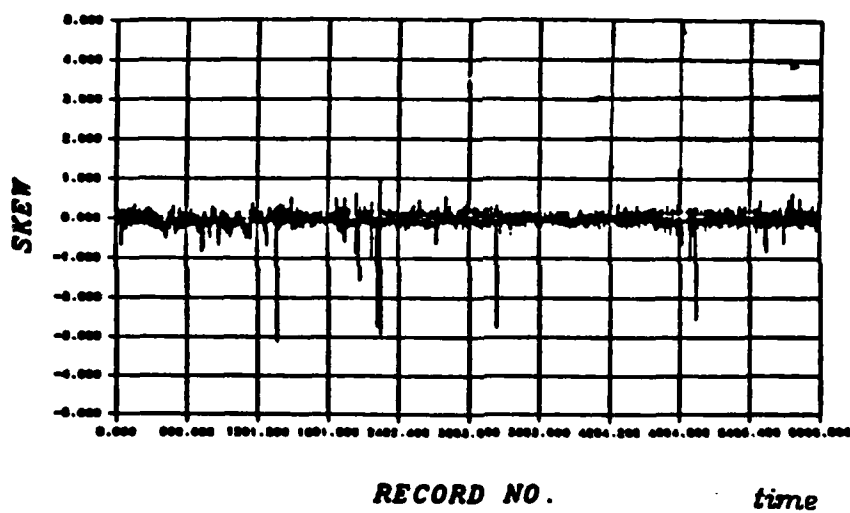
Fig. 3.I Statistical Moment of Under-Ice Noise (Skewness)

Fig. 3.J False-Alarm Rate vs. Threshold

(A Semi-Log Plot)

++++   Sample mean detector in Gaussian noise

_____   Sample mean detector under ice

▲▲▲▲   Modified sample mean detector under ice

Fig. 3.K Probability of Error vs. Threshold

(A Semi-Log Plot)

++++   Sample mean detector in Gaussian noise

_____   Sample mean detector under ice

▲ ▲ ▲ ▲   Modified sample mean detector under ice

3.8 Conclusions

The problem of detection with skewed noise is discussed. The sensitivity of the sample mean detector and the sign detector to the noise skewness is examined. When the underlying noise is Gaussian, perturbed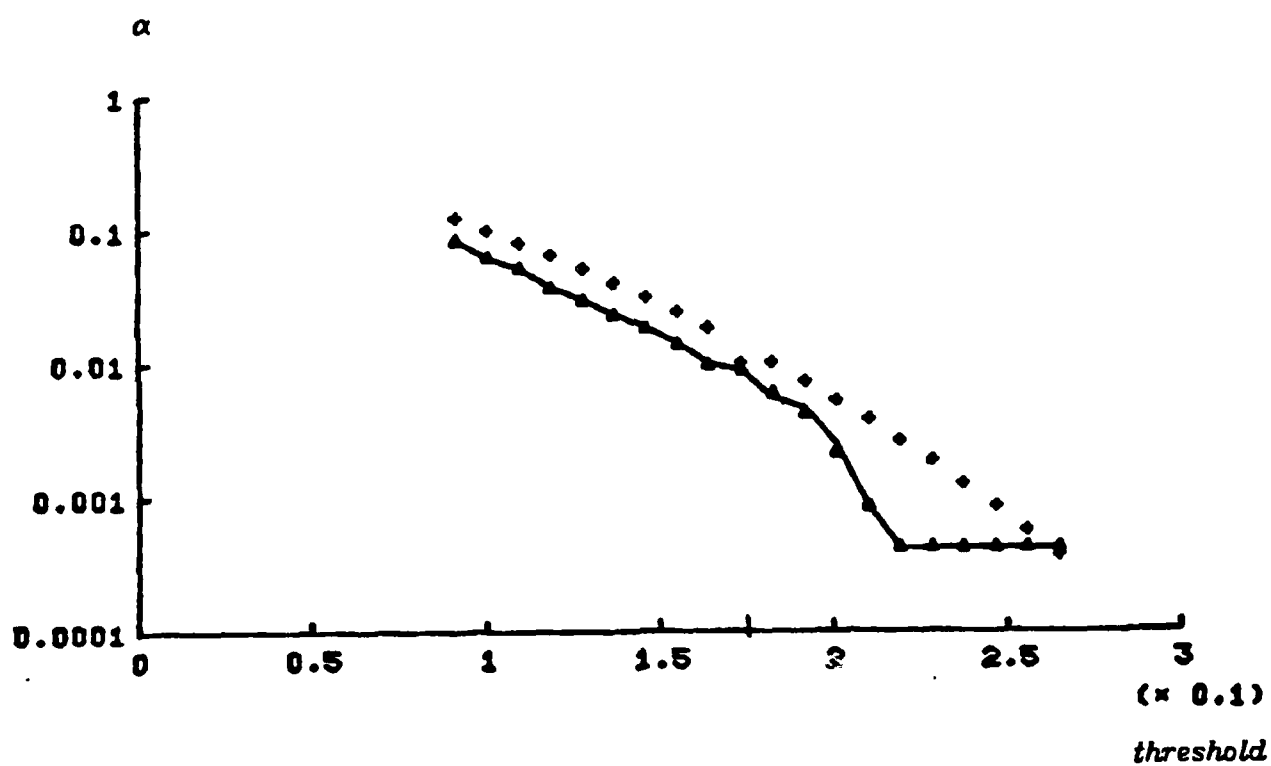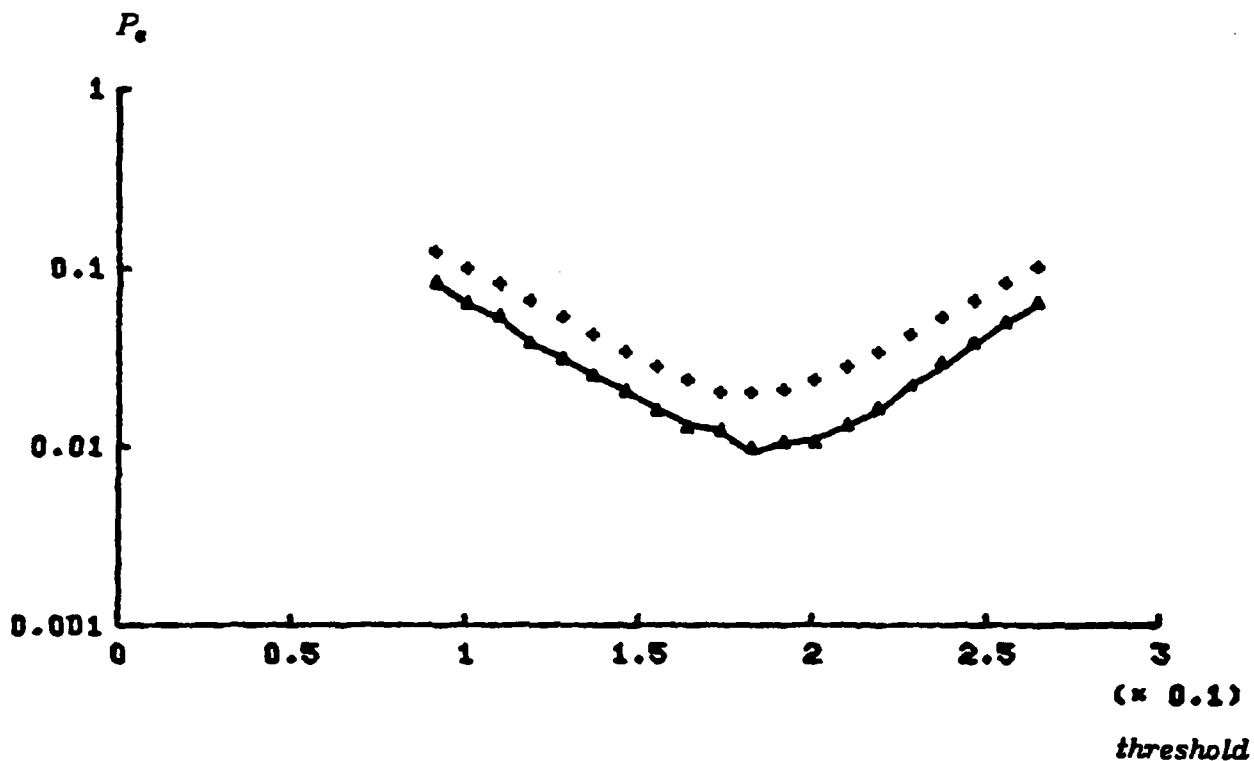 with a small amount of skewness, the sign detector is shown to be less sensitive than the sample mean detector. A modified sample mean detector is then proposed whose test statistic is constructed by adding a correction term to that of the sample mean detector. This correction term eliminates the lowest power of the skewness measure in the asymptotic expansion of the test statistic. Thus, when the skewness measure is small, this test statistic is virtually unaffected by noise skewness. Although the analysis here is based on the assumption of large sample size and small skewness measure, it does provide some general qualitative insight as to the skewness effect on detector performance. The simulation results in Section 3.6 provide a support to the analysis.

References

[1] J. Marcum, "Statistical Theory of Target Detection by Pulsed Radar," *IRE Trans. Information Theory, vol. IT-6, pp.59-267, 1960.*

[2] R.D. Martin and S.C. Schwartz, "Robust Detection of a Known Signal in Nearly Gaussian Noise," *IEEE Trans. Information Theory,* vol. IT-17, pp.50-56, 1971.

[3] S.A. Kassam and J.B. Thomas, "Asymptotically Robust Detection of a Known Signal in Contaminated Non-Gaussian Noise," *IEEE Trans. Information Theory,* vol. IT-22, pp.22-26. 1976.

[4] G.V. Trunk and S.F. George, "Detection of Targets in Non-Gaussian Sea Clutter," *IEEE Trans. Aerosp. Electron. Syst.,* vol. AES-6, pp.620-628, 1970.

[5] R.F. Dwyer, "Fram II Single Channel Ambient Noise Statistics," *101st Meeting of The Acoustical Society of America,* May, 1981.

[6] S.A. Kassam, G. Moustakides and J.G. Shin, "Robust Detection of Known Signals in Asymmetric Noise," *IEEE Trans. Information Theory,* vol. IT-28, pp 84-91, 1982

[7] N.J. Johnson, "Modified t-Test and Confidence Intervals for Asymmetrical Populations," *JASA,* vol. 73, pp.536-544, 1978.

[8] R.J. Carroll, "On Estimating Variances of Robust Estimators When The Errors Are Asymmetric," *JASA,* vol. 74, pp.674-679, 1979.

[9] M.G. Kendall and A. Stuart, The Advanced Theory of Statistics. vol-I New York: Hafner, 1963.

[10] W.G. Nichols and J.D. Gibbons, "Parameter Measure of Skewness," *Commun. Statist.,* vol.B-8, pp. 161-167, 1979.

[11] E.A. Cornish and R.A. Fisher, "Moments and Cumulants in The Specification of Distributions," Paper #31 in R.A. Fisher, Contributions to Mathematical Statistics. New York: John Wiley, 1950.

[12] D.L. Wallace, "Asymptotic Approximations to Distributions," *Ann. Math. Statist.,* vol. 29, pp. 635-654, 1958.

[13] A. Erdelyi, Asymptotic Expansions. New York: Dover, 1956.

[14] P. Billingsley, Probability and Measure. New York: John Wiley, 1979.

[15] J.W. Carlyle, "Nonparametric Methods in Detection Theory," Paper #1 in S.A. Kassam and J.B. Thomas, Nonparametric Detection: Theory and Applications. Dowden, Hutchinson and Ross. Stroudsburg, PA., 1981.

[16] J.B. Thomas, "Nonparametric Detection," *Proceedings of the IEEE*, vol. 58, pp.623-631, 1970.

[17] P.J. Bickel, "Edgeworth Expansions in Nonparametric Statistics," *Ann. Statist.*, vol.2, pp. 1-20, 1974.

CHAPTER 4

DETECTION OF SURE SIGNALS IN NOISE OF BOUNDED MAGNITUDE

4.1 Introduction

In the past few decades, a great amount of literature has been devoted to the theory of detection of signals in a noise environment [1]-[8]. To solve a detection problem, one usually has to impose some statistical assumptions on the noise and then employ methods of statistical inference. Needless to say, these statistical assumptions ought to be introduced with the consideration either of providing analytical tractability to the problem or of being compatible with practical applications. Unfortunately, these two factors are often contradictory to each other. For example, in many practical situations, the assumption that the noise sequence is a sequence of statistically well-defined random variables is not well justified. Hence the assumptions on the noise statistics will be just an approximation or simply a consequence of the desire for analytical tractability.

In most of the literature, the noise process considered usually has a probability density function with infinite support and the tail probabilities are extremely small, e.g. the Gaussian density function or the Laplace density function. Practically speaking, the tail area, considered sufficiently far out, is almost negligible. In fact, due to some environmental limitations, the noise may actually be bounded in magnitude. Furthermore, the availability of a well justified bound on noise magnitude is usually more realistic than that of the well justified statistical knowledge

of the noise. Thus the assumption of bounded noise may lead to more practical applicability of the solutions. As a matter of fact, bounded noise assumption has been employed in considering state estimaticn [9]-[10] and system identification problems [11].

This chapter considers a detection problem using a bounded noise assumption and proposes a rather different solution to this detection problem. To begin with, in Section 4.2, the signal is assumed to be a known positive constant and no statistical assumptions on the noise are employed. Then a sequential test procedure is devised based only on the knowledge of the noise bound. In contrast to the conventional hypothesis testing procedure, this procedure yields a singular solution. Specifically, it involves no possibility of making wrong decisions provided that the test procedure terminates. Now, one may argue that the presumed bound on the noise magnitude may not be precise; thus the resulting singular solution will not be realistic. One way to circumvent this problem is to introduce some randomness to the bound; namely, to suppose that the bound is a random variable with a known bounded distribution. Then the sequential procedure becomes a random test in the sense that the threshold and the associated performance are functions of a random variable. In Section 4.3, examples of some particular distributions of the noise sequence for a given realization of the presumed bound are used to demonstrate this example. Section 4.4 addresses the duality of hypothesis testing and estimation. Section 4.5 then discusses the problem of detecting an unknown signal in bounded noise where the detection problem will be formulated as a problem of estimating the signal.

4.2 Problem Formulation

Suppose that we consider the problem of deciding whether or not a constant signal is present by observing a noisy data sequence $\{x_i\}$. This problem can always be formulated as the following canonical hypothesis testing problem:

H: $x_i = n_i$

$$i = 1, 2, \ldots \tag{1}$$

K: $x_i = n_i + s$

The essential assumption to be discussed in this chapter is that the noise sequence $\{n_i\}$ is bounded in magnitude, namely

$$n_i^2 \leq B^2 \quad \text{for any } i, \text{ where } B > 0 \tag{2}$$

Now, if it is assumed here that the signal is a known positive constant, then

$$\text{under } H: \quad n_i^2 \leq B^2 \quad \Leftrightarrow \quad -B < x_i < B \tag{3}$$

$$\text{under } K: \quad n_i^2 \leq B^2 \quad \Leftrightarrow \quad s - B \leq x_i \leq s + B \tag{4}$$

From Fig. 4.A, it can be seen that $x_i < s - B$ only if H is true and $x_i > B$ only if K is true. Thus, to solve the detection problem, one can proceed as follows:

1. Obtain the data sequence $\{x_i\}$ sequentially.

2. For any i, if $x_i > B$, stop and decide K. On the other hand, if $x_i < s - B$, stop and decide H.

3. If $s - B < x_i < B$, obtain another observation x_{i+1} and repeat step 2.

It is clear now that if the bound B is known exactly, and if the test procedure terminates, one can expect no possibility of making wrong decisions. In contrast to conventional detection problems, this is a singular
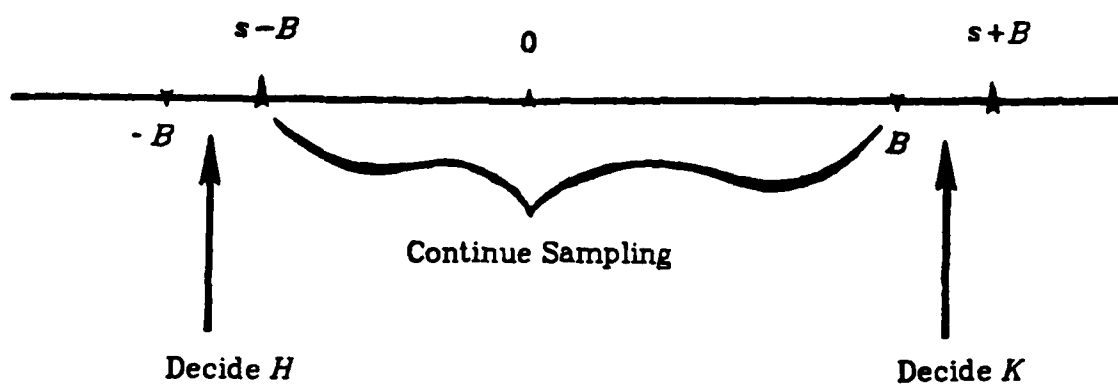
Fig. 4.A A Sequential Test Procedure

Assume constant bound;

$$0 \leq s \leq B$$

problem in the sense that this test procedure leads to a solution of zero false-alarm rate and unity power. This result is inherent from the assumption imposed here, namely, that the noise sequence is bounded in magnitude.

The sequential procedure discussed here involves some similarities to the standard sequential probability ratio test $(SPRT)$ [12]. In the $SPRT$, one assumes that the noise sequence is a sequence of random variables with a known probability distribution. Then the likelihood ratio L is computed and is usually a function of time. Two thresholds T_1 and T_2, with $0 < T_1 < 1 < T_2$, are chosen based on the desired values of the false-alarm rate and the power. The likelihood ratio L is then compared to the thresholds T_1 and T_2. If it exceeds T_2 at one of the comparison instants, the test terminates with the decision for K (signal present). If the likelihood ratio decreases below T_1, the test terminates with the decision for H (signal absent). If L remains between T_1 and T_2, another observation is made. On the other hand, instead of computing the likelihood ratio each time after making an observation, the sequential procedure discussed in this chapter simply compares the observed data x_i to the two thresholds B and $s - B$ to make decisions. Thus the test considered here requires less information about the noise and involves less computational complexity. However, another important criterion for evaluating performance of a sequential procedure is the number of samples, or the average sample number (ASN), required for termination. To consider the ASN, let us start by showing that the sequential procedure described here will terminate with probability one, under some general conditions.

Suppose that the probability that the observed data x_i lies in the interval $[s-B,B]$ is p_i, and that $0<p_i<1$ for all i. Furthermore, if the observation sequence $\{x_i\}$ is an independent sequence, then at the k-th observation, the probability that this test procedure has not terminated is $\prod\limits_{i=1}^{k} p_i$. However,

$$0 \le \lim_{k \to \infty} \prod_{i=1}^{k} p_i \le \lim_{k \to \infty} p_{\max}^k = 0$$

where

$$p_{\max} = \max_{1 \le i \le k} p_i$$

The above argument leads to the following lemma.

Lemma : Let $\{x_i\}$ be an independent sequence. Let $p_i = Prob.\ \{x_i \in [s-B,B]\}$. Then, if $0<p_i<1$ for all i, the sequential test procedure discussed in this chapter terminates with probability 1.

Knowing that the test procedure will terminate with probability 1, one can then consider the *ASN*. Now, if it is assumed that the observation sequence is independent and identically distributed, then, according to the strong law of large numbers [13], one may evaluate the *ASN* as follows:

(a). Suppose that the noise sequence $\{N_i\}$ is uniformly distributed between $[-B,B]$; then, under H:

$$Prob.\ \{(x:-B \le x \le s-B)\} = \frac{s}{2B}$$

Thus the expected value of the sample number M, or the *ASN*, is given by

$$E_H(M) = \frac{2B}{s} \quad \rightarrow \quad ASN$$

Similarly, under K:

$$E_K(M) = \frac{2B}{s} \quad \rightarrow \quad ASN$$

One can easily see that the two ASN's may not be equivalent under both hypothesis if one allows some asymmetry assumption. Furthermore, it should be noted that, if $B >> s$, then the probability that the observation will lie in $[-B, s-B]$ will be very small and thus the ASN will be very large. On the other hand, if B and s are comparable, one can expect a rapid termination of the test.

(b). If the noise sequence $\{N_i\}$ has a truncated Gaussian distribution, i.e. if it has a density function given by

$$f_N(n) \begin{cases} = \dfrac{1}{\sqrt{2\pi\sigma^2}(1-2\Phi(-B/\sigma))} \exp^{-n^2/2\sigma^2} & \text{if } |n| \le B \\ = 0 & \text{otherwise} \end{cases}$$

Then it can be shown similarly that

$$ASN = \frac{(1 - 2\Phi(-B/\sigma))}{[\Phi(\frac{s-B}{\sigma}) - \Phi(-B/\sigma)]}$$

under either H or K.

4.3 Constant Known Signals in Noise of Random Bounds

In the previous section, the problem of detecting a known signal in a magnitude bounded noise has been discussed. A sequential test procedure was proposed which leads to a singular solution. However, there are some obvious problems associated with this result. The presumed bound may not be tight, and thus the number of samples required to ter-

minate the test procedure may be unnecessarily large. Furthermore, in practice, the bound may not be known exactly, and could even be assumed smaller than it actually is. Therefore, to make the problem more realistic, one may assume that the presumed bound is a random variable with a known distribution function. Then, clearly, the two thresholds will also be random variables and the solution thus obtained will involve non-zero probabilities of false-alarm and miss.

Now, for simplicity of discussion, one may suppose that the presumed bound B is a random variable uniformly distributed between $[r-\varepsilon, r+\varepsilon]$, where r is the actual bound of the noise sequence $\{n_i\}$. A realization of B will be denoted by b. Then, assuming that the signal is a known positive constant, the two thresholds $s-B$ and B will be random variables uniformly distributed between $[s-r-\varepsilon, s-r+\varepsilon]$ and $[r-\varepsilon, r+\varepsilon]$, respectively. It can be seen, Fig. 4.B, that a false-alarm (type I error) occurs when $b > r$ and $x_i > r$. Hence the probability of false-alarm (false-alarm rate) is given by

$$\alpha(b) = Prob.\{b > r, x_i > r\}$$
$$= Prob.\{x_i > r \mid b > r\} \, Prob.\{b > r\} \qquad (7)$$

Moreover, a miss (type II error) occurs when $b > r$ and $x_i < s-r$. Hence the probability of miss is

$$1-\beta(b) = Prob.\{b > r, x_i < s-r\}$$
$$= Prob.\{x_i < s-r \mid b > r\} \, Prob.\{b > r\} \qquad (8)$$

Thus the power $\beta(b)$ is given by

$$\beta(b) = 1 - Prob.\{x_i < s-r \mid b > r\} \, Prob.\{b > r\} \qquad (9)$$

It is seen in Eq. (7) and Eq. (8) that both the false-alarm rate α and
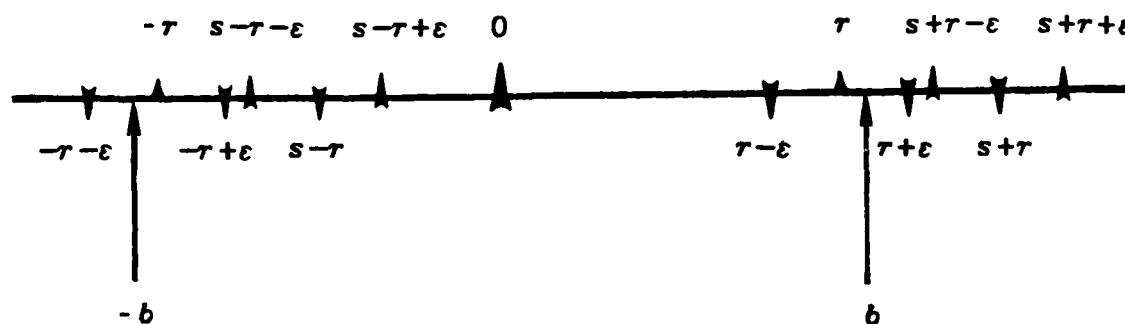
A False-alarm occurs when $b > r$ and $x_i \in [\, r, b \,]$

A Miss occurs when $b > r$ and $x_i \in [\, s - b, s - r \,]$

Fig. 4.B A Sequential Test Procedure with Random

Bounds On Noise Magnitude

the power β are functions of a random variable B and thus are also random variables. Thus this test procedure can be regarded as a random sequential procedure as its thresholds are random and also its performance. Properties of this test procedure will clearly depend on the statistical distributions of the false-alarm rate and the power. The following sections employ some statistical assumptions on the noise sequence and provide clearer insights.

4.3.1 Uniformly Distributed Random Noise

Suppose that for any presumed bound b, the noise sequence $\{n_i\}$ is uniformly distributed between $[-b,b]$. Thus, from Eq. (1), under H, $\{x_i\}$ is uniformly distributed on $[-b,b]$ and, under K, $\{x_i\}$ is uniformly distributed on $[s-b,s+b]$. It may be interesting to see that the probability density function of X_i under H is given by

$$
p_X(x)
\begin{cases}
= \dfrac{1}{4\varepsilon}\ln(\dfrac{\tau+\varepsilon}{x}) & \text{if } \tau+\varepsilon \geq x > \tau-\varepsilon \\[2mm]
= \dfrac{1}{4\varepsilon}\ln(\dfrac{\tau+\varepsilon}{\tau-\varepsilon}) & \text{if } \tau-\varepsilon \geq x > -\tau+\varepsilon \\[2mm]
= \dfrac{1}{4\varepsilon}\ln(\dfrac{\tau+\varepsilon}{-x}) & \text{if } -\tau+\varepsilon \geq x > -\tau-\varepsilon \\[2mm]
= 0 & \text{otherwise}
\end{cases}
\tag{10}
$$

Also the probability density function of X_i under K is simply a shift to the right of Eq. (10). As discussed previously, false-alarms occur when $b>\tau$ and $x_i>\tau$, while misses occur when $b>\tau$ and $x_i<s-\tau$. The false-alarm rate and the power can then be evaluated via Eqs. (7) and (9), respectively. From Eq. (7),

$$
\alpha(b) = \text{Prob.}\{x_i>\tau \,|\, b>\tau\} \, \text{Prob.}\{b>\tau\}
\tag{11}
$$
$$
= \frac{b-\tau}{2b} \cdot \frac{1}{2} = \frac{1}{4} - \frac{\tau}{4b}
$$

Similarly, from Eq. (9),

$$\beta(b) = Prob.\{x_i < s - r \mid b > r\} \, Prob.\{b > r\}$$
$$= \frac{3}{4} + \frac{r}{4b} \tag{12}$$

Now that both the false-alarm rate and the power are functions of a random variable [Eqs. (11) and (12)] it is essential to consider their statistical properties; e.g. the probability distributions and the statistical moments. The probability density function of $\alpha(b)$ is given by

$$f_A(\alpha) \begin{cases} = \dfrac{2r}{\varepsilon(1-4\alpha)^2} & \text{if } \dfrac{-\varepsilon}{4(r-\varepsilon)} \le \alpha \le \dfrac{\varepsilon}{4(r+\varepsilon)} \\ = 0 & \text{otherwise} \end{cases} \tag{13}$$

and that of $\beta(b)$ by

$$f_B(\beta) \begin{cases} = \dfrac{2r}{\varepsilon(1-4\beta)^2} & \text{if } \dfrac{3}{4} + \dfrac{r}{4(r+\varepsilon)} \le \beta \le \dfrac{3}{4} + \dfrac{r}{4(r-\varepsilon)} \\ = 0 & \text{otherwise} \end{cases} \tag{14}$$

Moreover, the first two statistical moments of the false-alarm rate and the power are given by

$$E[\alpha(b)] = \frac{1}{8} - \frac{1}{8t} ln(1+t) \tag{15}$$

where

$$t = \frac{\varepsilon}{r}$$

$$E[\beta(b)] = \frac{7}{8} + \frac{1}{8t} ln(1+t) \tag{16}$$

and

$$var[\alpha(b)] = var[\beta(b)]$$
$$= \frac{1}{64} + \frac{1}{32(1+t)} - \frac{1}{32t} ln(1+t) - \frac{1}{64t^2}[ln(1+t)]^2 \tag{17}$$

Eqs. (13) and (14) show that both α and β have bounded distributions with an algebraic tail behavior. In practice, one may be interested in cases

where $t \ll 1$, then

$$\ln(1+t) = t - \frac{1}{2}t^2 + \frac{1}{3}t^3$$
$$\approx t - \frac{1}{2}t^2$$

Thus

$$E[\alpha(b)] \approx \frac{t}{16}$$
$$E[\beta(b)] \approx 1 - \frac{1}{16}t$$

and

$$Var[\alpha(b)] = Var[\beta(b)] \approx \frac{1}{32}t^2$$

Hence if $t \ll 1$, the first-order moments of both α and β are monotone linear functions of t while the variance is a quadratic function of t. This result suggests a reasonable conclusion that, the smaller t is, the smaller is the probability of error involved in this test procedure. Figs. 4.C-4.E depict the first two moments of α and β as functions of t as t ranges from 0.001 to 0.1 .

4.3.2 Random Noise With Triangular Distribution

An example of a unimodal and bounded distribution is the triangular distribution whose density function is given by

$$f_X(x) \begin{cases} = \frac{1}{b}(1 - \frac{|x|}{b}) & \text{if } |x| \leq b \\ = 0 & \text{otherwise} \end{cases}$$

Again, by assuming that the presumed bound B is uniformly distributed between $[r-\varepsilon, r+\varepsilon]$, one may evaluate the false-alarm rate and the power via Eqs. (7) and (9), respectively.

$$\alpha(b) = Prob. \{x_i > r \mid b > r\} Prob. \{b > r\}$$

$$= \left[\int_r^b \frac{1}{b}(1 - \frac{|x|}{b})dx \right] \times \frac{1}{2}$$

$$= \left(\frac{b-r}{2b} \right)^2$$

Now, the probability density functions of $\alpha(b)$ and $\beta(b)$ are given by

$$f_A(\alpha) \begin{cases} = \frac{1}{4\varepsilon} \frac{r}{\sqrt{\alpha}(1-2\sqrt{\alpha})^2} + \frac{1}{4\varepsilon} \frac{r}{\sqrt{\alpha}(1+2\sqrt{\alpha})^2} & \text{if } \frac{\varepsilon^2}{4(r-\varepsilon)^2} \geq \alpha \geq \frac{\varepsilon^2}{4(r+\varepsilon)^2} \\ = 0 & \text{otherwise} \end{cases}$$

and

$$f_B(\beta) \begin{cases} = \frac{1}{4\varepsilon} \frac{r}{\sqrt{1-\beta}(1-2\sqrt{1-\beta})} + \frac{1}{4\varepsilon} \frac{r}{\sqrt{1-\beta}(1+2\sqrt{1-\beta})} & \text{if } 1 - \frac{\varepsilon^2}{4(r-\varepsilon)^2} \geq \beta \geq 1 - \frac{\varepsilon^2}{4(r+\varepsilon)} \\ = 0 & \text{otherwise} \end{cases}$$

The first two statistical moments of α and β are given by

$$E[\alpha(b)] = \frac{1}{8} + \frac{1}{8(1+t)} - \frac{1}{4t}ln(1+t)$$

$$\approx \frac{t^2}{8} \qquad \text{if } t \ll 1 \tag{18}$$

and

$$E[\beta(b)] = \frac{7}{\varepsilon} - \frac{1}{8(1+t)} + \frac{1}{4t}ln(1+t)$$

$$\approx 1 - \frac{t^2}{8} \qquad \text{if } t \ll 1 \tag{19}$$

$$Var[\alpha(b)] = Var[\beta(b)]$$

$$= \frac{1}{64} - \frac{5}{96t} - \frac{1}{16t}ln(1+t) - \frac{1}{16t^2}[ln(1+t)]^2 + \frac{5}{32(1+t)}$$

$$+ \frac{1}{16} \frac{1}{t(1+t)}ln(1+t) - \frac{1}{64} \frac{1}{(1+t)^2} + \frac{1}{16t(1+t)^2} - \frac{1}{96t(1+t)^3}$$

$$\approx \frac{13}{96}t^2 \qquad \text{if } t \ll 1 \tag{20}$$

A similar set of curves to those in Figs. 4.C-4.E describing the first two moments of α and β as functions of t are given in Figs. 4.F-4.H. Here, instead of an approximately linear relation, an approximately quadratic relation between the mean and t is shown. This relation can also be seen
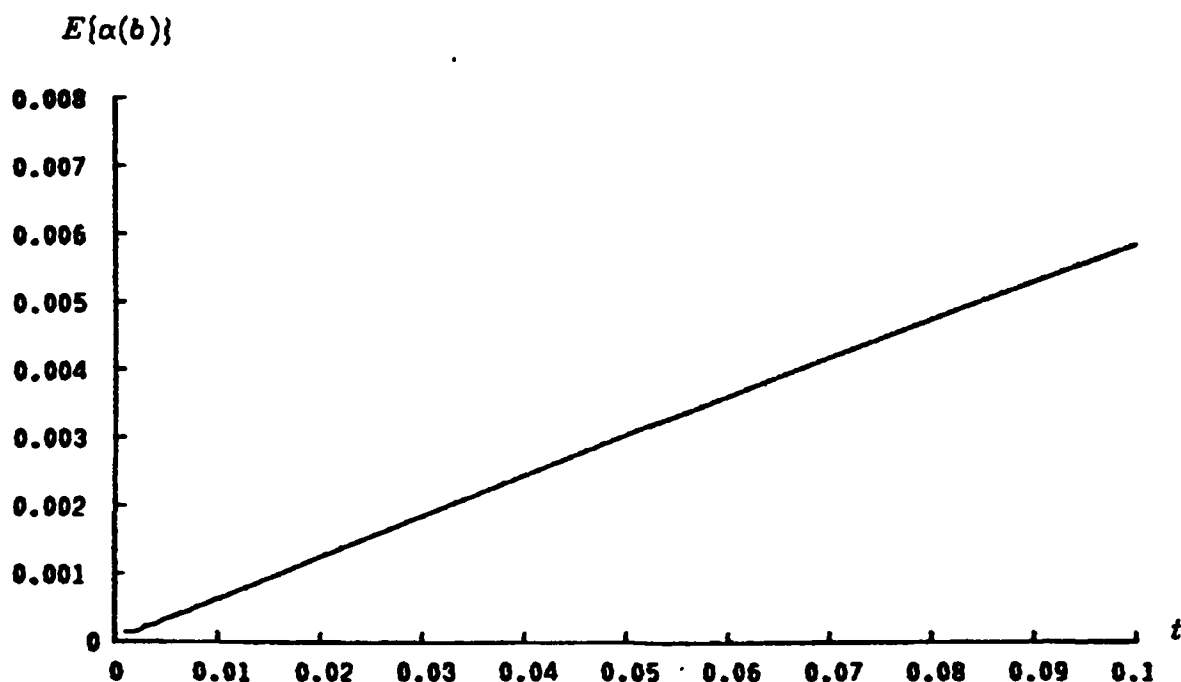
Fig. 4.C Mean of α

Noise is uniformly distributed in [-b, b].

$$t \triangleq \frac{\varepsilon}{\gamma}$$

Fig. 4.D Mean of β

Noise is uniformly distributed in [-b, b].

$$t \triangleq \frac{\varepsilon}{\gamma}$$

$$Var[\alpha(b)] = Var[\beta(b)]$$

Fig. 4.E Variance of α and β
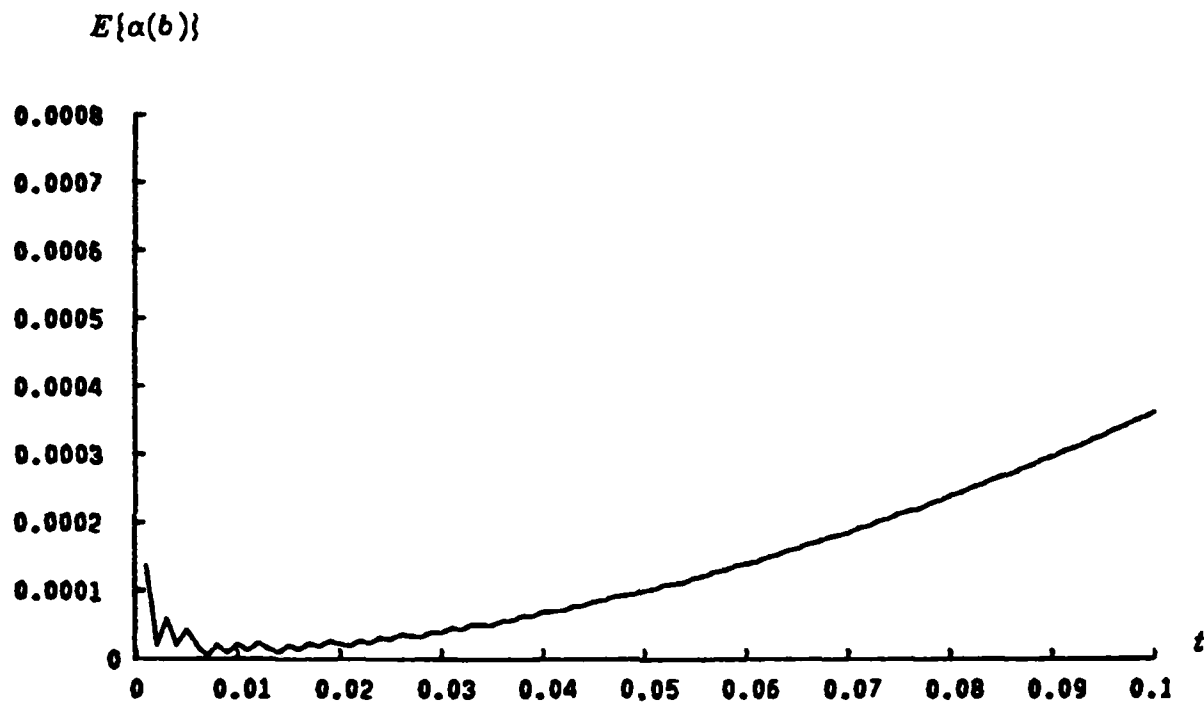
Noise is uniformly distributed in [-b, b].

$$t \triangleq \frac{\varepsilon}{\gamma}$$

$E\{\alpha(b)\}$

Fig. 4.F Mean of α

Noise is triangularly distributed in [-b, b].

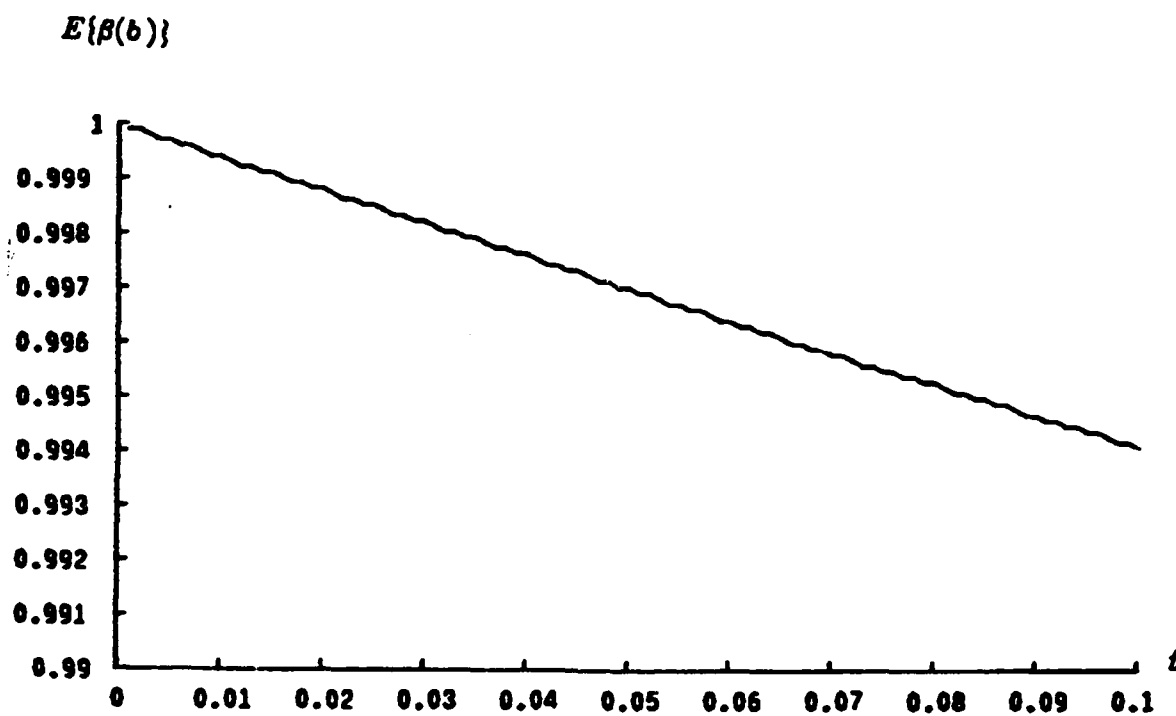$$t \triangleq \frac{\varepsilon}{\gamma}$$

$E\{\beta(b)\}$

Fig. 4.G Mean of β

Noise is triangularly distributed in [-b, b].

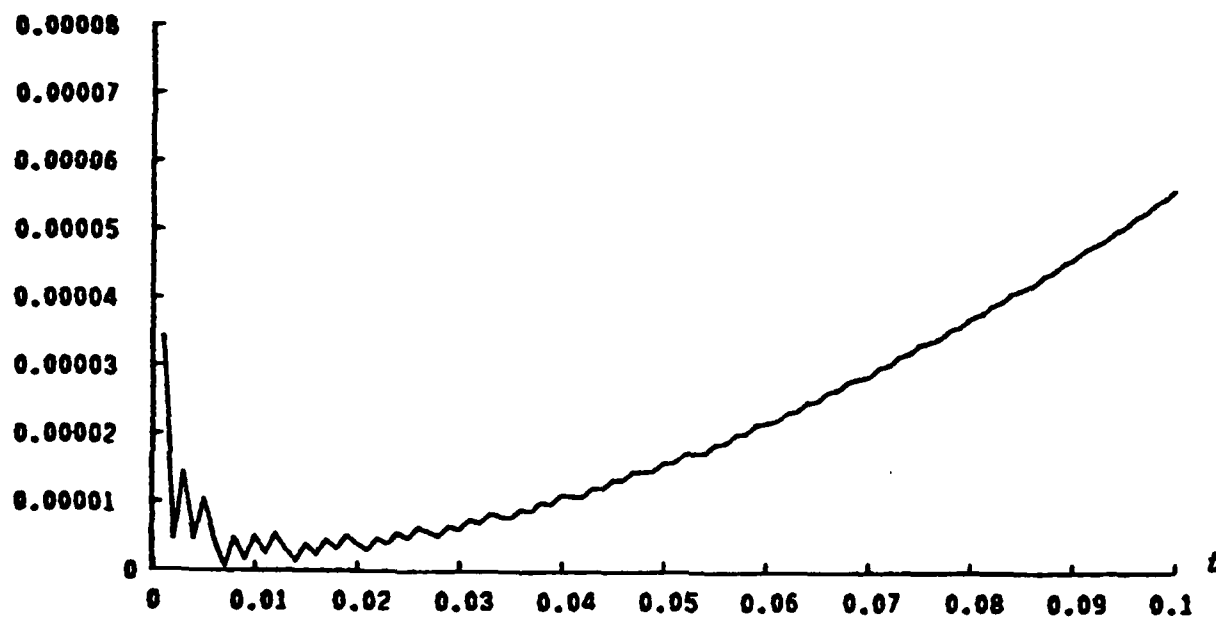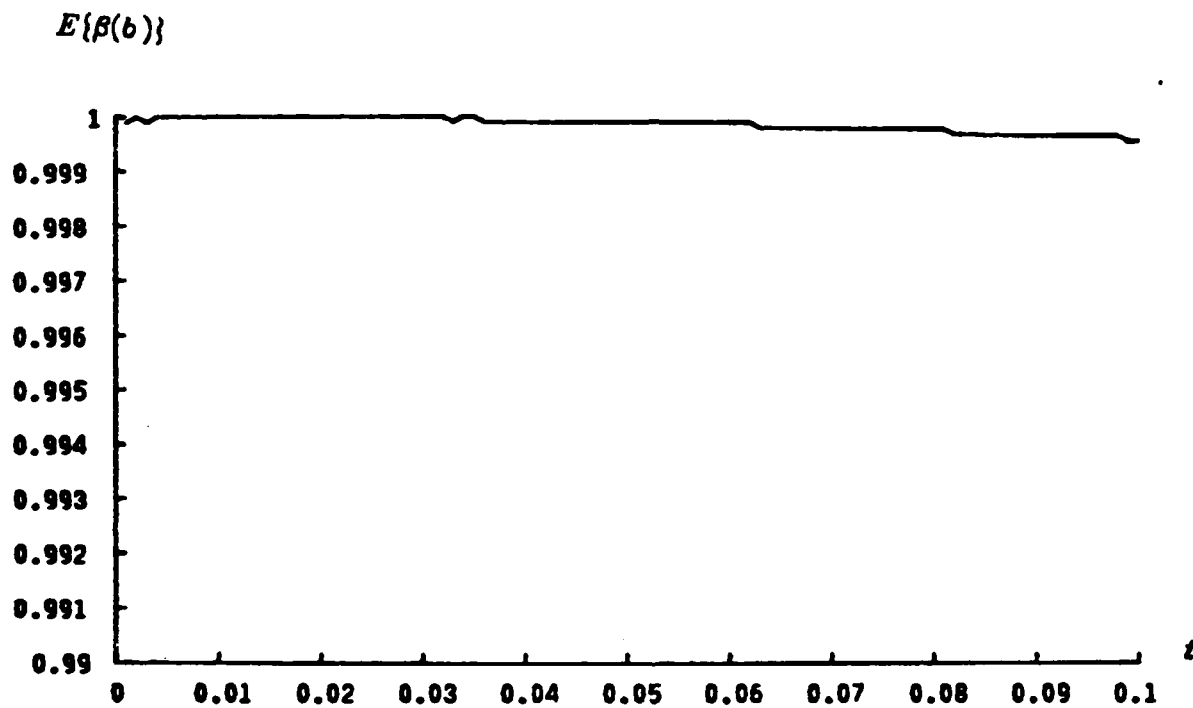$$t \triangleq \frac{\varepsilon}{\gamma}$$

$Var[\alpha(b)] = Var[\beta(b)]$

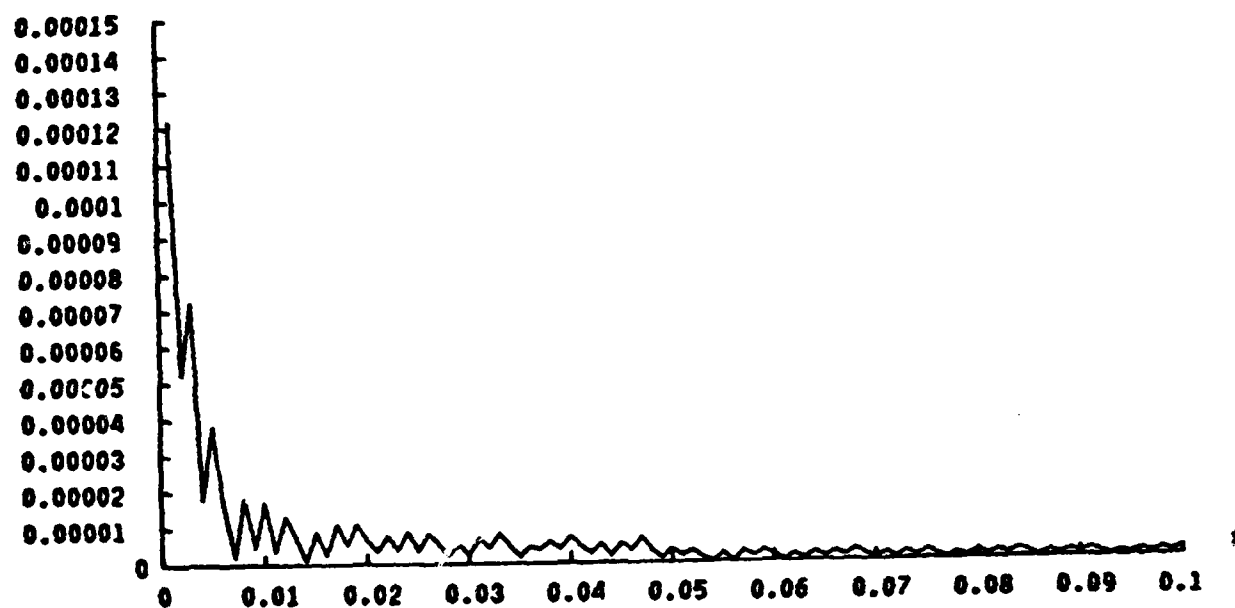Fig. 4 H Variance of α and β

Noise is triangularly distributed in [-b, b].

$$t \triangleq \frac{\varepsilon}{\gamma}$$

from Eqs. (18) and (19). The difference certainly results from the different assumptions on the noise distributions employed. However, as can be seen from Figs. 4.F-4.H, the effect of inaccuracy of the presumed bound on the noise on the probability of error is less than that in the previous case. This is due to a smaller tail probability of the noise distribution used here.

4.3.3 Random Noise with Truncated Gaussian Distribution

In the last two sections, the noise sequence was assumed to have a bounded distribution, i.e. uniform distribution and triangular distribution, with a random bound which has a uniform distribution. However, these distributions may not be of sufficient practical interest. In applications such as radar or sonar, the noise density function often exhibits a Gaussian-shaped central part, although the magnitude of the noise may be bounded by some equipment limitations. The probability density function of a random variable with a truncated Gaussian distribution is given by

$$
f_X(x) \begin{cases} = \dfrac{C(b)}{\sqrt{2\pi}\sigma} \exp^{-x^2/2\sigma^2} & \text{if } |x| \le b \\ = 0 & \text{otherwise} \end{cases} \tag{21}
$$

where $C(b)$ is a normalization constant and is given by

$$
C(b) = [\Phi(\tfrac{b}{\sigma}) - \Phi(\tfrac{-b}{\sigma})]^{-1}
$$

As in the previous sections, the bound b is assumed to be a random variable uniformly distributed between $[r-\varepsilon, r+\varepsilon]$. Then, by Eqs. (7) and (9), the false-alarm rate and power are given by

$$\alpha(b) = Prob. \{x_i > r, b > r\}$$
$$= \frac{C(b)}{2}[\Phi(\frac{b}{\sigma}) - \Phi(\frac{r}{\sigma})]$$

and

$$\beta(b) = 1 - \frac{C(b)}{2}[\Phi(\frac{s-r}{\sigma}) - \Phi(\frac{s-b}{\sigma})]$$

Because of the functional forms involved in the density function, Eq. (21), the statistical moments of α and β are hardly expressible in a simple form. However, curves depicting them as functions of t can be obtained by using numerical methods as shown in Figs. 4.I-4.K. Again, these curves are obtained for t varying from 0.001 to 0.1 . Also, the value of r is set at 4σ, as the tail probability of a Gaussian distribution beyond this point is sufficiently small. A smooth monotone nondecreasing relation between t and the first two moments of the probability of error can be seen from these curves. This is compatible with the intuition that the more accurate the presumed bound is, the less probability of error will result from this test procedure.

To investigate in more detail the sequential test procedure here, one may be interested in comparing its performance to that of the standard *SPRT*. It is obvious that both error probabilities converge to zero as the bound on the noise distribution goes to infinity, namely

$$\lim_{r \to \infty} \alpha(b) = 0 \qquad \text{almost surely}$$

and

$$\lim_{r \to \infty} \beta(b) = 1 \qquad \text{almost surely}$$

Unfortunately, the *ASN* will be infinite in the limit. A similar result can be expected from the standard *SPRT*. Thus, to compare these two test

Fig. 4.I Mean of α

The noise has a truncated Gaussian distribution.

$$t \triangleq \frac{\varepsilon}{\gamma}$$

$$\gamma = 4\sigma$$

END
FILMED
DTIC
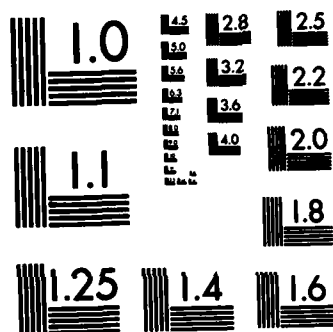
MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A
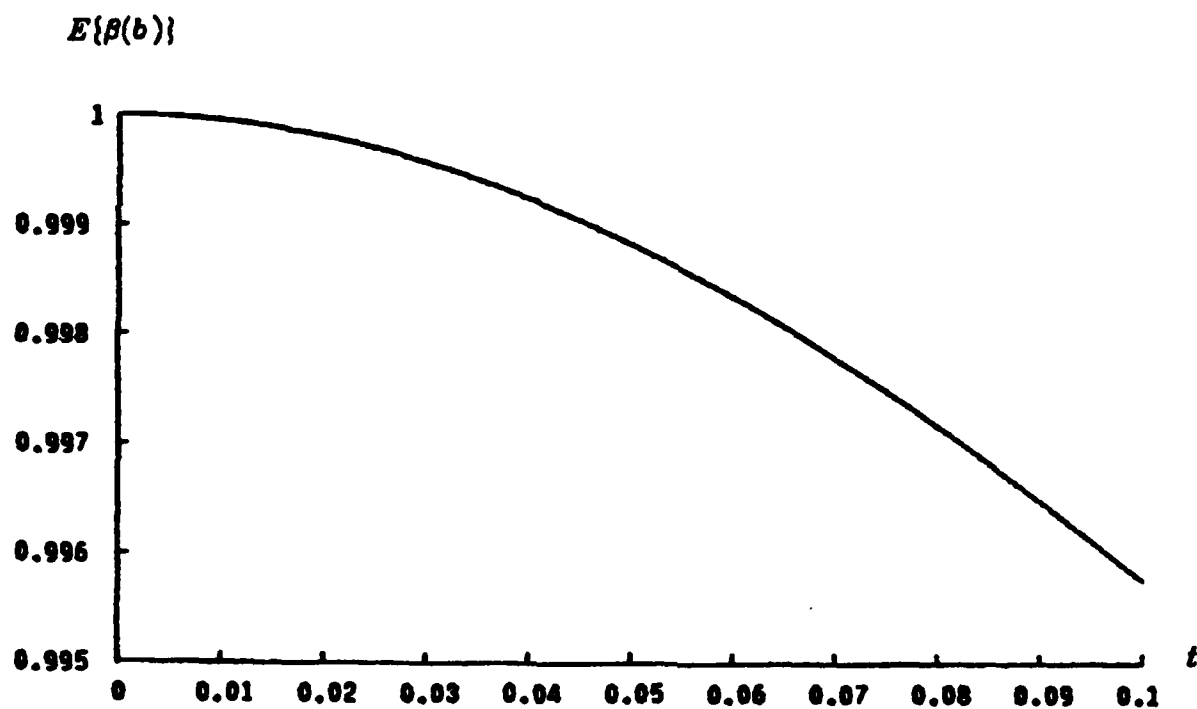
Fig. 4.J Mean of β

The noise has a truncated Gaussian distribution.

$$t \triangleq \frac{\varepsilon}{\gamma}$$

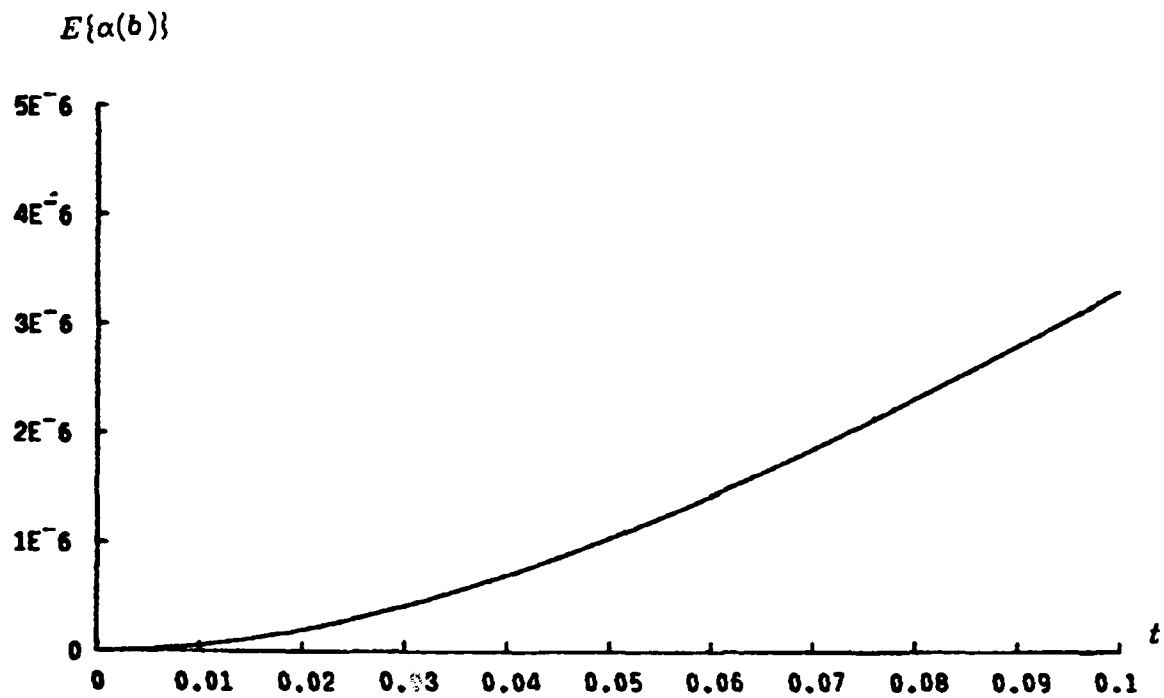$$\gamma = 4\sigma, \quad \frac{s^2}{\sigma^2} = 6.25$$
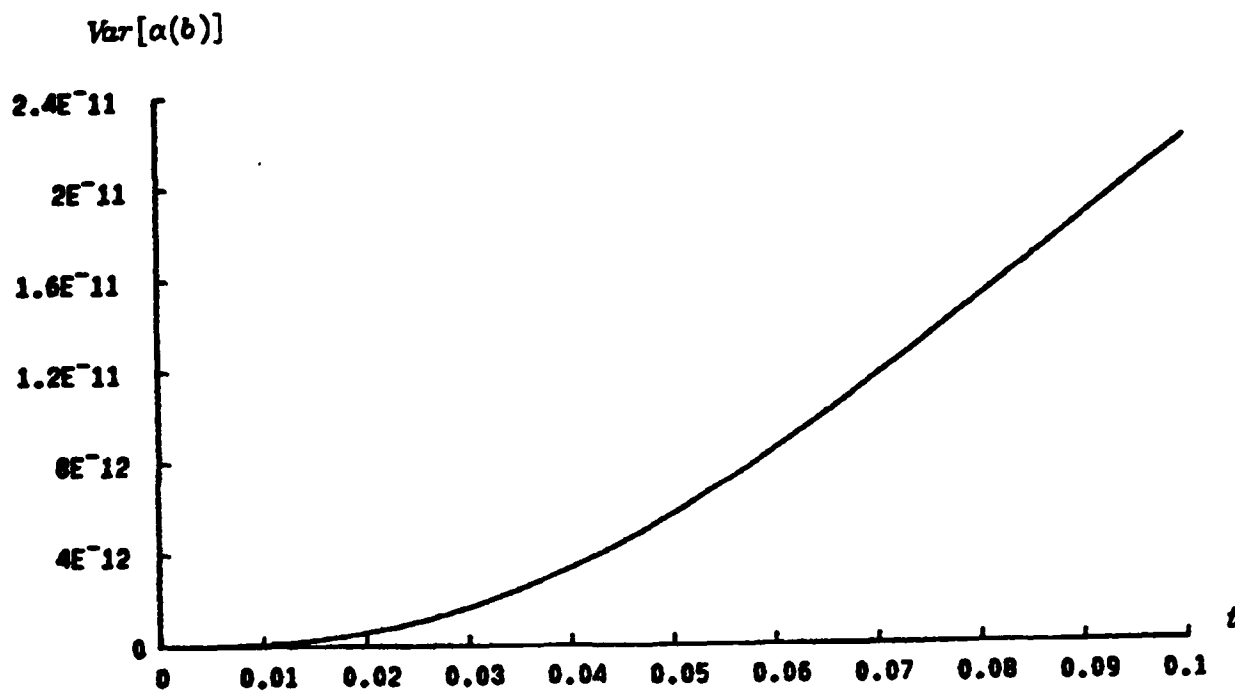
Fig. 4.K Variance of α

The noise has a truncated Gaussian distribution.

$$t \triangleq \frac{\varepsilon}{\gamma}$$

$$\gamma = 4\sigma$$

procedures, one has to consider rate of convergence of the error probabilities. In [12], it is shown that the *ASN* of the *SPRT* under H and under K are given respectively by

$$E_H(M) = \frac{(1-\alpha) \log T_1 + \alpha \log T_2}{E_H(z)}$$

and

$$E_K(M) = \frac{\beta \log T_1 + (1-\beta) \log T_2}{E_K(z)}$$

where z is the test statistic and T_1 and T_2 are two thresholds given by

$$T_1 = \beta / (1-\alpha) \quad \text{and} \quad T_2 = (1-\beta) / \alpha$$

Thus one can see that for the same α and the same β, the *ASN* of the *SPRT* is less than that of the sequential procedure here. On the other hand, a complete knowledge on the noise density function required by the *SPRT* and the computational simplicity involved in the proposed sequential procedure should also be taken into account in making comparisons.

4.3.4 Random Noise with Truncated Laplace Distribution

Another distribution which could be of practical interest is the Laplace distribution. As in the previous section, the truncated version of the Laplace distribution will be considered here with probability density function given by

$$f_X(x) \begin{cases} = \dfrac{hC(b)}{2} \exp^{-h|x|} & \text{if } |x| \le b \\ = 0 & \text{otherwise} \end{cases}$$

where $C(b)=(1-e^{-hb})^{-1}$ is a normalization constant. Again, by assuming that the bound b is a uniformly distributed random variable on $[\tau-\varepsilon, \tau+\varepsilon]$, one may evaluate the false-alarm rate and the power via Eqs.

(7) and (9),

$$\alpha(b) = Prob. \{z_i > r, b > r\}$$
$$= \frac{C(b)}{4} [e^{-hr} - e^{-hb}]$$

and

$$\beta(b) = 1 - \frac{C(b)}{4} [e^{h(s-r)} - e^{h(s-b)}] \quad \text{assuming that } s > r$$

Again, in order to achieve some insights as to the statistical properties of α and β, one has to employ numerical techniques. Figs. 4.K-4.N. shows the first two moments of α and β as functions of t. A smooth monotone nondecreasing relation between t and the probability is also shown here. However, comparing to the results in the previous section, greater probability of error is obtained here due to the heavier tail probability in the Laplace distribution.

4.4 Duality Between Hypothesis Testings and Confidence Regions

In the previous section, detection of a known signal in noise has been formulated as a canonical hypothesis testing problem, namely,

$$H: z_i = n_i$$

$$i = 1, 2, \ldots$$

$$K: z_i = n_i + s, \quad s > 0$$

This is equivalent to testing $H: \vartheta = \mu_0$ vs $K: \vartheta = \mu_0 + s$, where $\vartheta = E\{X_i\}$ and $\mu_0 = E\{N_i\}$. However, if the signal strength is unknown, this is equivalent to testing $H: \vartheta = \mu_0$ vs $K: \vartheta \neq \mu_0$; i.e. any value of ϑ other then μ_0 is a possible alternative. One can then consider this problem as one of estimating the mean of the observation data sequence. `s a matter of fact, there exists a complete duality between families of level α tests and level $(1-\alpha)$

Fig. 4.L Mean of α

The noise has a truncated Laplace distribution, $h = \sqrt{2}$.

$$t \triangleq \frac{\varepsilon}{\gamma}$$

$$\gamma = 4\sigma$$

$E\{\beta(b)\}$

Fig. 4.M   Mean of β

The noise has a truncated Laplace distribution, $h = \sqrt{2}$.

$$t \triangleq \frac{\varepsilon}{\gamma}$$

$$\gamma = 4\sigma; \quad \frac{s^2}{\sigma^2} = 6.25$$
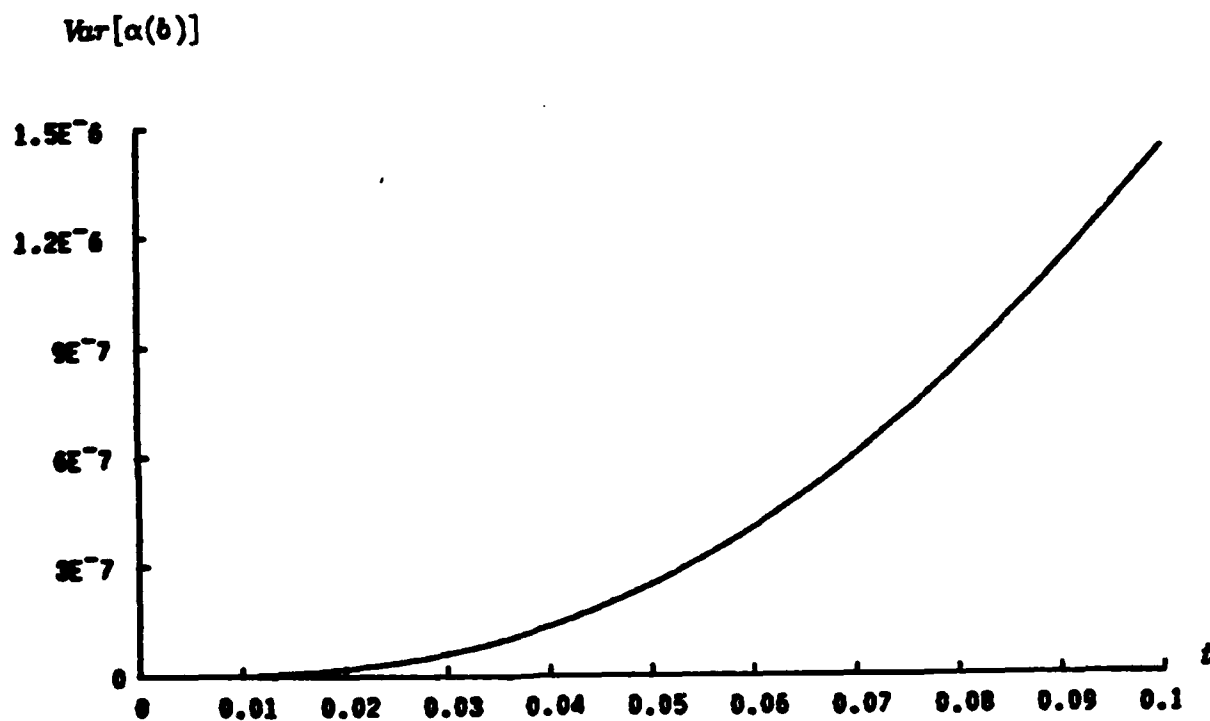
Fig. 4.N Variance of α
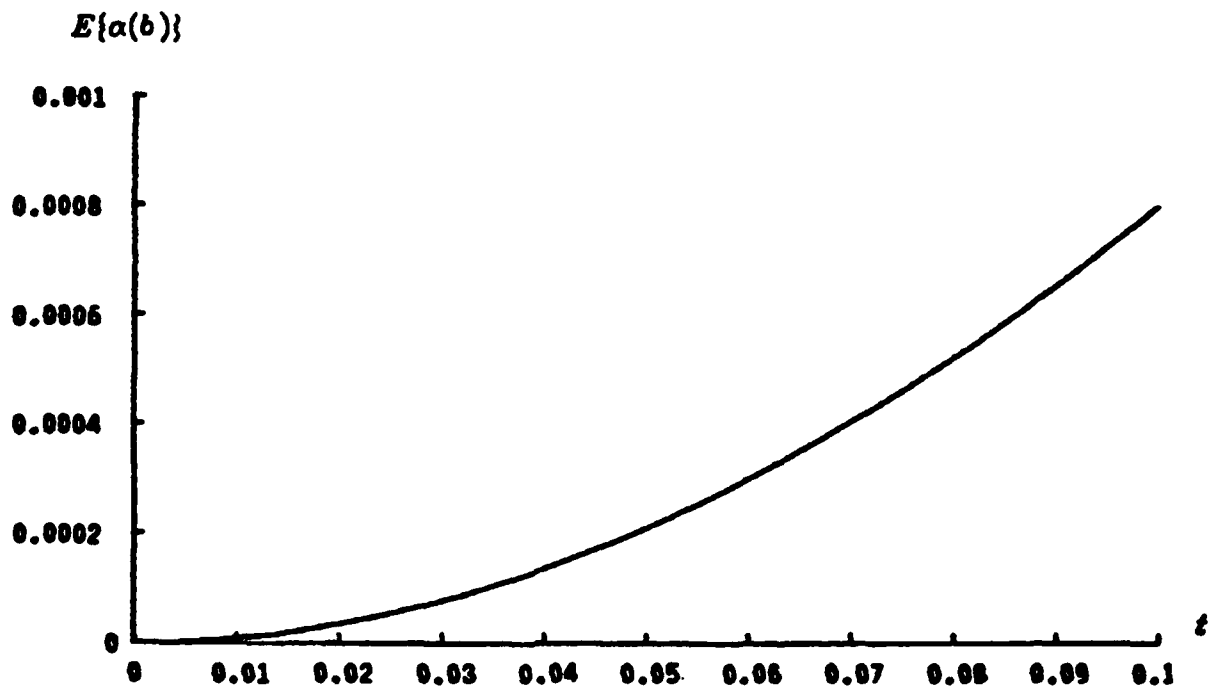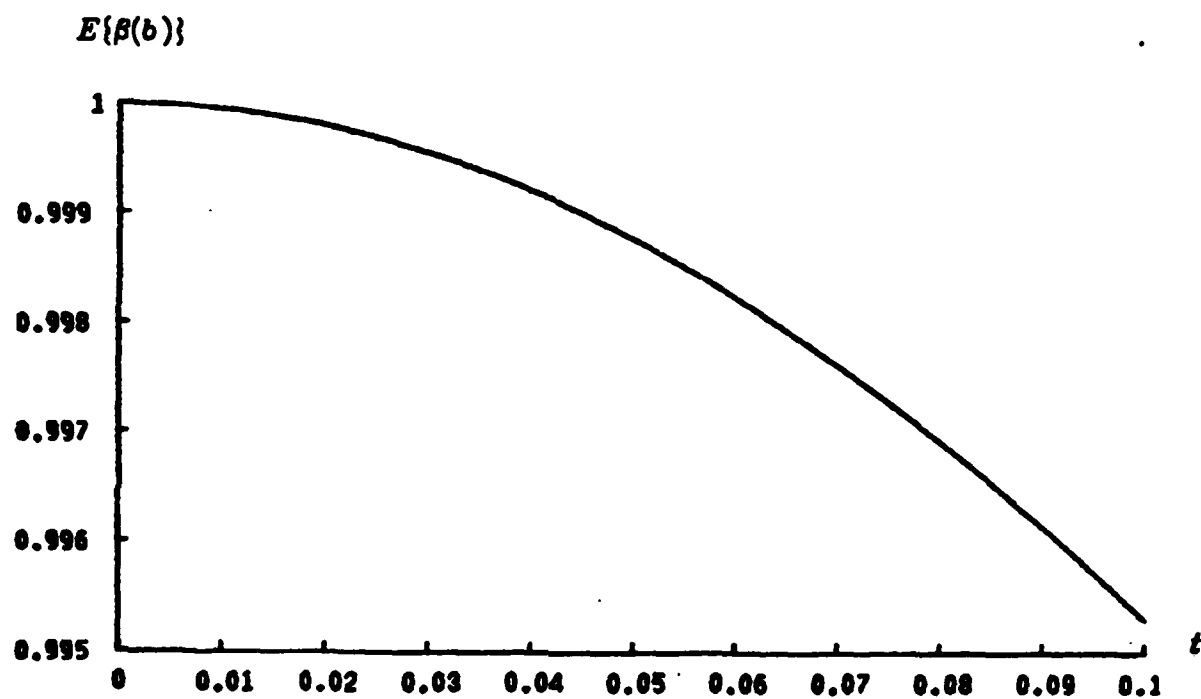
The noise has a truncated Laplace distribution, $h = \sqrt{2}$.

$$t \triangleq \frac{\varepsilon}{\gamma}$$

$$\gamma = 4\sigma$$

confidence regions. To explain this concept in more detail, let us consider the following example [14]: Suppose that, on the basis of m observations, $x_1, x_2, ..., x_m$, one wants to decide, with the assumption that the x_i's are i.i.d., whether or not the population mean of $\{X_i\}$ is μ_0. Then a size-α test can be based on a level $(1-\alpha)$ confidence interval as follows: Accept H, if and only if, the postulated value μ_0 is a member of the level $(1-\alpha)$ confidence interval $[\bar{X} - a, \bar{X} + a]$ where a is a function of $(1-\alpha)$ and μ_0. Thus, if a likelihood ratio test is considered, by properly choosing the threshold T_0, one can associate the test with the confidence interval. However, since the same interval is used for every μ_0, there actually is generated a family of level α tests with parameter ϑ. Conversely, families of tests can generate confidence bounds and intervals. Let $\{\delta(X, \vartheta)\}$ be a family of tests such that $\delta(X, \mu_0)$ is a test of level of significance α for testing $H: \vartheta = \mu_0$ for each $\mu_0 \in \Theta$ where $\Theta \subset R$ and where $\delta(X, \vartheta)$ denotes the critical function of a test. For fixed x, define the subset $C(x)$ of Θ by

$$C(x) = \{\vartheta : \delta(x, \vartheta) = 0\}$$

This is just the set of all ϑ that would be accepted if $X = x$ is observed and the given family of tests is used. Suppose that $C(x)$ is of the form $<a(x), \infty) \cap \Theta$ for each x, where $<$ indicates that the point $a(x)$ may or may not be in the indicated ray. Then, $a(x)$ is a lower confidence bound of level $(1-\alpha)$ for ϑ. This is true since

$$Prob. \{a(X) \leq \vartheta\} \geq Prob. \{\vartheta \in C(x)\} = Prob. \{\delta(X, \vartheta) = 0\}$$
$$= 1 - Prob. \{\delta(x, \vartheta) = 1\} \geq 1 - \alpha$$

Similarly, if $C(x)$ is of the form $<a(x), b(x)> \cap \Theta$ for each x, then $[a(x), b(x)]$ is a level $(1-\alpha)$ confidence interval for ϑ. Also, if $C(x)$ is of

— 99 —

the form $(-\infty, b(x)> \cap \Theta$ for each x, then $b(x)$ is an upper confidence bound of level $(1 - \alpha)$ for ϑ. In fact, both confidence bounds and intervals can be considered as random subsets of the parameter space. The true parameter value is included in these subsets with probability at least $(1-\alpha)$.

The above discussion demonstrated a complete duality between confidence intervals and hypothesis testings and also provides the motivation for casting the hypothesis testing problem as an estimation problem.

4.5 Detection of Unknown Constant Signals in Bounded Noise

The problem of detecting known constant signals in bounded noise has been discussed in the previous parts of this chapter. When the bound on the noise quantity is a known constant, a sequential procedure which leads to a trivial solution to the detection problem was considered. On the other hand, when the bound is a random variable, this sequential procedure becomes a random test. By imposing some statistical assumption, e.g. known distribution, on the bound and also on the noise, one can then describe some statistical properties of this random test. In particular, the first two moments, or even the probability density functions, of the false-alarm rate and the power can be obtained.

In this section, a similar problem will be considered. However, the signal strength will be assumed to be unknown. To solve this detection problem, one may consider either of the following two approaches: (1) Formulate performance of the test procedure as a function of the signal

strength and then optimize, in some sense, the performance with respect to the signal strength. (2) Cast the detection problem as an estimation problem. In the last section, duality between hypothesis testings and confidence regions has been discussed. It was shown that the detection of an unknown constant signal can be considered as an estimation of the mean of the observation data.

Now the hypothesis testing problem given in Eq. (1) is considered, where s is unknown and $n_i^2 \leq B^2$, for any i. As discussed in Section 4.4, this problem can be formulated as follows:

$$H_1 : \vartheta = 0$$
$$K_1 : \vartheta \neq 0$$

where $\vartheta = E\{X_i\}$ and, of course, we have assumed that $E\{N_i\} = 0$ for any i. One of the easiest ways to estimate the mean of the observation data is to consider the sample mean, $S_M = \frac{1}{M} \sum_{i=1}^{M} X_i$. However, we now proceed rather differently as follows:

Let $G_i = \{z : x_i - B \leq z \leq x_i + B\}$ and take intersections of G_i's for $i = 1, 2, ...,$ and define

$$I_M = \bigcap_{i=1}^{M} G_i$$

It is clear that the true value of the mean of the x_i's should lie inside I_M, for each M. In this section, we consider the estimate of the mean to be a set rather than a single point. Thus every point in I_M is an estimate here. If a set function μ_M is defined for the measure of I_M as follows:

$$\mu_M = l(I_M) = \text{length of } I_M,$$

then, obviously, μ_M is a monotone nonincreasing function of M. Now, if

we assume that $G_i - G_j \neq \phi$ for any $i \neq j$, then

$$\lim_{M \to \infty} \mu_M = 0,$$

and thus the true value of the mean evolves in the limit. Therefore, based on these arguments, one can solve the detection problem as follows:

1. If $x_i^2 \geq B^2$, stop and decide K

2. If $x_i^2 \leq B^2$, obtain G_i and I_i.

3. When $\mu_i \leq \varepsilon$ stop. If $I_i \supset \{0\}$, decide H, otherwise decide K. Here ε is a small constant which is properly chosen so that the sequential procedure can be truncated at some finite number of samples.

The only possibility of making wrong decisions in this procedure is to decide H while K is true, i.e., the false-alarm. The signal strength and the choice of ε clearly affect this possibility. Again, the assumption of constant bound on the noise quantity leads to a singular solution. Another reason for this result is that the sequential procedure discussed here does not utilize any statistical information, which in some practical situations may not be easy to obtain, and thus does not involve any statistical procedure. One disadvantage associated with this sequential procedure is that the number of samples required to terminate the test may be too large. However, an obvious advantage here is the computational simplicity involved in implementing this procedure.

4.6 Conclusions

In this chapter, detection of a constant signal in bounded noise has been considered. Both the known signal and unknown signal cases are addressed. In the former, if the bound is a known constant, the solution is singular, namely, both the false-alarm rate and the probability of miss are zero. To make the problem more realistic, we considered noise with a random bound; then a random test is obtained. Some statistical properties of the performance parameters, e.g. the false-alarm rate and the power, are obtained for several examples with presumed assumptions on the distributions of the noise and the bound. As can be expected, the probability of error of the test is dependent on the tail property of the noise distribution and, of course, the accuracy of the presumed bound on the noise magnitude. In contrast to the *SPRT*, the test procedure discussed in this chapter requires very little statistical information on the noise and involves less computational complexity, although it may be terminated with a larger sample size. Finally, detection of an unknown signal in bounded noise is considered as an estimation problem. The test procedure discussed in Section 4.5 is rather different from the usual approaches in the sense that, at each time, a set of estimates instead of a single estimate is obtained. Unfortunately, performance evaluation of this procedure is not clear. More complete results may be obtained in the future.

References

[1]. J. Marcum, "Statistical Theory of Target Detection by Pulsed Radar", *IRE Trans. on Information Theory*, Vol. IT-6, pp. 59-267, 1960.

[2]. J. Capon, "Locally Optimum Nonparameters Methods for the Detection of Signals in Noise," *Proc. of the 1963 IEEE Symposium on Adaptive Processes*, pp. 64-77, 1963.

[3]. M. Kanefsky and J.B. Thomas, "On Adaptive Nonparametric Detection Systems Using Dependent Samples", *IEEE Trans. Inform. Theory*, Vol. IT-11, pp. 521-526, 1965.

[4]. J.W. Carlyle and J.B. Thomas, "On Nonparametric Signal Detectors", *IEEE Trans. Inform. Theory*, Vol. IT-10, pp. 146-152, 1964.

[5]. H.L. Van Trees, Detection, Estimation and Modulation Theory, Part I, New York:John Wiley, 1968.

[6]. C.W. Helstrom, Statistical Theory of Signal Detection, New York:Pergamon Press, 1968.

[7]. R.D. Martin and S.C. Schwartz, "Robust Detection of a Known Signal in Nearly Gaussian Noise," *IEEE Trans. Inform. Theory*, Vol. IT-17, pp. 50-56,1971.

[8]. A.H. El-Saway and V.D. Vandelinde, "Robust Sequential Detection of Signals in Noise," *IEEE Trans. Inform. Theory*, Vol. IT-25, pp. 346-353, 1979.

[9]. F.C. Schweppe, "Recursive State Estimation: Unknown but Bounded Errors and System Inputs," *IEEE Trans. Automat. Control.* Vol. AC-13, pp. 22-28, 1968.

[10].F.M Schlaepfer and F.C. Schweppe, "Continuous-Time State Estimation Under Disturbances Bonded by Convex Sets," *IEEE Trans. on Automat. Contr.*, Vol. AC-17, pp 197-205, 1970.

[11].E. Fogel and Y.F. Huang, "On The Value of Information in System Identification- Bounded Noise Case," *Automatica* Vol. 18, pp 229-238, 1982.

[12].A. Wald, Sequential Analysis, New York:Dover, 1973.

[13].W. Feller, An Introduction to Probability Theory and Its Applications, Vol. I, Third Edition, New York:John Wiley & So: , 1968.

[14].P.J. Bickel and K.A. Doksum, Mathematical Statistics, San Francisco:Holden Day, 1977.

CHAPTER 5

A SET-THEORETIC APPROACH TO DETECTION PROBLEMS

5.1 Introduction

In the literature, detection problems have almost always been discussed by the formulations of statistical inference. Statistical models are used to describe the noise or the noise and the signal. These formulations usually require some presumed statistical information which may not be available precisely. This chapter considers the detection problem from a rather different point of view. Set-theoretic formulations, instead of statistical formulations, are employed to describe the input model. This approach may be useful in detecting the failure of dynamical systems, where problems usually occur when the components of the system deviate from the nominal ones [1] - [3].

The problem addressed here can be considered as a generalization of the one discussed in Section 4.5 of the last chapter, where detection of an unknown signal in bounded noise is considered. It is assumed here that the signal is an unknown constant vector. The noise sequence is constrained by a compact set in R^k. The detection problem is solved by estimating the signal first. Then, based on the estimate of the signal, decisions are made as to whether or not the signal is present. The procedure can be explained briefly as follows: at each observation, a set which is compatible with the constraint and the observation data is obtained. The estimate of the signal based on each single observation is thus a set rather than a single element in the k-dimensional vector space. When multiple observations are taken, intersection of these sets

is considered as the set of estimates. Then, under some conditions, this set will eventually converge to a point which is the signal to be estimated. The decision of whether a signal is absent clearly depends on whether this point is located at the origin.

5.2 A Detection Problem

Given an observation sequence $\{x_i\}$, one is required to decide, subject to some constraint, whether or not an unknown constant vector-signal is present. This problem can be formulated as the usual hypothesis testing problem as follows:

$$H : x_i = n_i$$
$$K : x_i = s + n_i \qquad i = 1, 2, 3, \ldots\ldots \qquad (1)$$

where x, s, and n are k-vectors. Also n is a random sequence constrained by the following relation

$$n_i \in S_i \qquad (2)$$

where S_i is a compact set in R^k. In accordance with the discussion in Sections 4.4 and 4.5 of the previous chapter, this detection problem will be cast as an estimation problem here. The estimates of the signal discussed here will be a set rather than a single vector in the k-dimensional vector space. Thus the solution of the detection problem of Eq.(1) is equivalent to finding a set Ω_i in R^k with Ω_i being compatible with Eqs.(1) and (2). Note that every element in Ω_i is an estimate of the signal s.

To find the set Ω_i is conceptually straightforward. From each observation, a set compatible with Eqs.(1) and (2) is obtained. Every element in this set is thus an estimate based on this observation. After i

observations are made, the set of estimates is simply the intersection of these sets obtained at each observation. Hence the set Ω_i can be expressed as

$$\Omega_i = \bigcap_{j=1}^{i} S_j \qquad (3)$$

Obviously, the sequence of sets $\{\Omega_i\}$ is a monotone non-increasing sequence, namely

$$\Omega_1 \supseteq \Omega_2 \supseteq \Omega_3 \supseteq \dots\dots$$

Under some specific assumptions, this sequence will converge to a point. The decision of whether or not the signal is present then depends on where this point is located in the R^k. Similar to the results obtained in the previous chapter, the solution to this detection problem is singular due to the set-theoretic assumptions used here.

Unfortunately, in general, formulations of the sets Ω_i are not computationally simple. Suppose that, for example, the constraint set S_i is an intersection of two half spaces. Then at each time instant, the set Ω_i defined in Eq.(3) is a polytope. Formulation of a polytope is usually a numerically complicated task. One possibility to avoid this difficulty is to define a set which can be more easily formulated and which bounds Ω_i "tightly" (of course, the word "tightly" has to be precisely defined). A good candidate for the bounding set in this case will be a k-dimensional ellipsoid which can be formulated in terms of quadratic equations. Note, however, that when the bounding set is used to proceed with the estimation procedure, there may exit a non-empty subset of this bounding set that contains elements which are not compatible with the constraints and the observed data. Thus it is no longer true that every element of the set

is an estimate. Furthermore, when only a finite number of observations is used, decisions based on this set of estimate may not be correct. Hence false-alarms and misses will occur.

5.3 Detection of Constant Signals in Bounded k-Variate Noise

The detection problem given in Eq.(1) is considered here again. Now, the noise sequence is subject to the following constraint

$$\mathbf{n}_i^T \mathbf{n}_i \leq 1 \qquad \forall i \qquad (4)$$

Under H, Eq. (4) implies that

$$\mathbf{x}_i^T \mathbf{x}_i \leq 1 \qquad (5)$$

Under K, it becomes

$$(\mathbf{x}_i - \mathbf{s})^T (\mathbf{x}_i - \mathbf{s}) \leq 1 \qquad (6)$$

After some algebraic manipulations, Eq. (6) becomes

$$(\mathbf{s} - \mathbf{s}_c)^T P_i^{-1} (\mathbf{s} - \mathbf{s}_c) \leq 1 \qquad (7)$$

where

$$\mathbf{s}_c = P_i \mathbf{x}_i \qquad P_i = I_{k \times k}$$

The Eq. (7) stands for a k-dimensional ellipsoid centered at $\mathbf{s}_c$. In fact, it is a k-dimensional spheroid since P_i is an identity matrix. If one takes multiple observations, each observation would result in a k-dimensional ellipsoid as given by Eq. (7). Thus, after taking M observations, the set of estimates of the signal is the intersection of all these ellipsoids, and the geometric center should be a good estimate of $\mathbf{s}$, when a single estimate is required. However, the intersection of ellipsoids is very difficult to formulate; hence its geometric center becomes a fuzzy concept. A possible way to get around this problem is to consider an

ellipsoid which "tightly" bounds the intersection. A "tightly bounding ellipsoid" is defined as follows [4]

Definition : An ellipsoid E which bounds a set S is said to be tight if $E \supseteq E' \supseteq S$ implies that $E' = E$, for any subset E' of E.

Nevertheless, this definition only defines a "tightly" bounding ellipsoid in a certain direction. There are infinitely many "tightly" bounding ellipsoids. In order to obtain an efficient procedure, these ellipsoids should be chosen in some optimal way.

Let $\{E_i\}$ be a sequence of optimal bounding ellipsoids and S_i be the sets specified by Eq. (6), then,

$$E_{i-1} = \{ s : (s - s_c(i-1))^T P_{i-1}^{-1} (s - s_c(i-1)) \le 1 \} \tag{8}$$

and

$$S_i = \{ s : (x_i - s)^T (x_i - s) \le 1 \}$$

An ellipsoid which "tightly" bounds the intersection of E_i and S_i is formulated as follows

$$E = \{ s : (s - s_c)^T P_{i-1}^{-1} (s - s_c) + \rho_i (x_i - s)^T (x_i - s) \le 1 + \rho_i \}$$

where $0 \le \rho_i < \infty$. After some algebraic manipulations, one can show that

$$E = \{ s : (s - s_c)^T Q^{-1} (s - s_c) \le \delta \} \tag{9}$$

where

$$Q^{-1} = P_{i-1}^{-1} + \rho_i I_{k \times k}$$
$$s_c = Q [P_{i-1}^{-1} s_c(i-1) + \rho_i x_i]$$

and

$$\delta = 1 + \rho_i - (s_c(i-1) - x_i)^T P_{i-1}^{-1} Q (s_c(i-1) - x_i)$$

Clearly, the parameter ρ_i determines the orientation (and even the

size) of the bounding ellipsoid. Finding a sequence of optimal bounding ellipsoids $\{E_i\}$ is equivalent to finding a sequence of optimal $\{\rho_i\}$. One can, for example, consider an ellipsoid with minimum volume as an optimal one. It is known that the volume of an ellipsoid formulated by Eq. (9) is linearly proportional to the determinant of the matrix Q^{-1}/δ. Thus minimizing the determinant with respect to ρ_i will result in an optimal bounding ellipsoid. The problem of finding optimal bounding ellipsoids with minimum volumes has been examined in the literature [5] and thus will not be repeated here. It should be noted here that both E_{i-1} and S_i contain the true signal vector. Therefore the intersection of them is not an empty set. Furthermore, it can be shown that δ is a positive real number; thus the ellipsoid described by Eq. (9) is non-degenerate. Fig. 5.A shows the bounding ellipsoid which contains the intersection of the observation set and the previous ellipsoid.

Now, a sequential algorithm for solving the detection problem of Eq. (1) can be developed in the following:

(i) Start with a very large spheroid centered at the origin, namely

$$E_0 = \{ s : s^T P_0^{-1} s \leq 1 \}$$

where

$$P_0^{-1} = \sigma I_n$$

and σ is a very large real number.

(ii) Successively take observations, compute $z_i^T z_i$ for every i. If one ever obtains
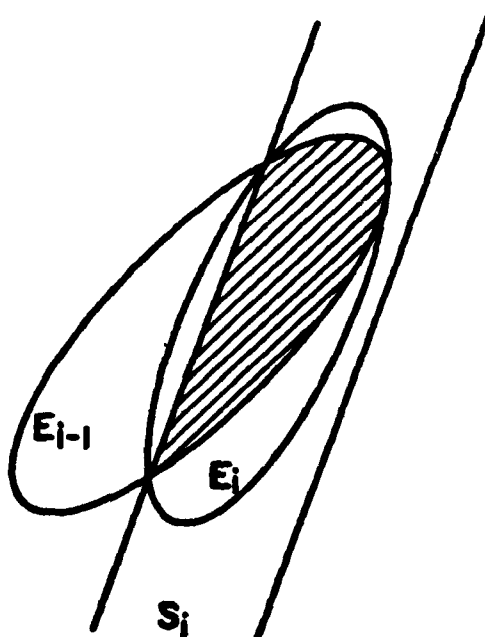
$$z_i^T z_i > 1 \qquad \text{for any } i$$

Fig. 5.A   An Illustration of The Bounding Ellipsoid E_i
and The Observation Set S_i

decide K. Otherwise, calculate an optimal bounding ellipsoid based on the current observation and the previous ellipsoid.

(iii) As the number of observation increases, $s_c(i)$ should become closer and closer to s. Therefore, decide H if $s_c(i) \to 0$, otherwise, decide K.

However, use of the decision criteria, Step (iii), may require an infinite number of samples. In reality, the procedure has to be stopped at some finite number of samples. To make the algorithm more practical, one may choose a neighborhood of the origin O such that if $s_c(i) \in O$ and if $\Omega_i \supset \{0\}$, H will be accepted.

5.4 Performance Evaluation

Since the sequential estimation-detection algorithm discussed in this chapter is significantly different from the conventional ones, its performance will also be evaluated from a rather different point of view. Notice that the set E_i defined in Eq. (8) is different from the set of estimates Ω_i as specified in Eq. (3). As a matter of fact, it is usually true that the difference set of E_i and Ω_i is a non-empty set. Let $D_i \triangleq E_i - \Omega_i$, then the elements of D_i are not estimates that are compatible with the constraint set and the observation data. To evaluate the performance of this detection algorithm, we proceed as follows :

Let $\mu_D(i)$ and $\mu_E(i)$ be the set functions for the measures of D_i and E_i respectively. Obviously,

$$\lim_{i \to \infty} \mu_D(i) = 0 \qquad \text{and} \qquad \lim_{i \to \infty} \mu_E(i) = 0$$

Based on the sequential algorithm here, the false-alarm occurs when $E_i \supset \{0\}$ but $s_c(i) \in O$ at the stop. Also, the miss occurs when $E_i \supset \{0\}$ and $s_c(i) \in O$ at the stop. To be more specific, one should define a distance measure $d(s_1, s_2)$ (which may be the Euclidean norm) of two vectors s_1 and s_2 on the space R^k. Then a neighborhood of the origin O is defined by properly choosing a small positive real number ε such that

$$O = \{s : d(s,0) < \varepsilon\}$$

The decision criteria now is re-stated below

    (a) If $x_i^T x_i > 1$ for some i, decide K.

    (b) If $E_i \not\supset \{0\}$, decide K.

    (c) If $E_i \supset \{0\}$ while $d(s_c(i),0) > \varepsilon$, decide K. On the other hand, if $E_i \supset \{0\}$ and $d(s_c(i),0) < \varepsilon$, decide H.

5.5 Concluding Remarks

A detection problem using no statistical assumptions on the input model has been discussed in this chapter. The only assumption imposed here is that the noise is constrained by a compact set in R^k. Due to this set-theoretic assumption, the solution thus obtained is rather different. The detection scheme devised here is essentially a combination of estimation and detection with a sequential nature. However, since it is required to estimate the signal, the discrimination between H, the hypothesis, and K, the alternative, may be insignificant when a weak signal is encountered.

As for evaluating the performance of this scheme, set functions are

introduced. Unfortunately, due to the computational difficulties involved in formulating intersections of compact sets, the result presented here is somewhat preliminary. Finally, it should be stressed here that finding a sequence of optimal bounding ellipsoids, Section 5.3, yields a facility for evaluating the information, contained in each observation set, which is pertinent to the updating of the estimates. This reduces the computational complexity involved in the estimation procedure.

References

[1] A.S. Willsky and H.L. Jones, "A Generalized Likelihood Ratio Approach to The Detection and Estimation of Jumps in Linear Systems," *IEEE Trans. on Automat. Contr.*, Vol. AC-21, pp.108-112, 1976.

[2] A.S. Willsky, "A Survey of Design Methods for Failure Detection in Dynamical Systems," *Automatica*, Vol. 12, pp. 601-611, 1976.

[3] T.H. Kerr, "False Alarm and Correct Detection Probabilities over a Time Interval for Restricted Classes of Failure Detection Algorithms," *IEEE Trans. Inform. Theory*, Vol. IT-28, pp. 619-631, 1982.

[4] W. Kahan, "Circumscribing An Ellipsoid about An Intersection of Two Ellipsoids," *Can. Math. Bull.*, Vol. 11, No.3, pp.437-441, 1968.

[5] E. Fogel and Y.F. Huang, "On The Value of Information in System Identification-Bounded Noise Case," *Automatica*, Vol. 18, pp. 229-238, 1982.

CHAPTER 6

CONCLUSIONS

6.1 Summary of Results

The method of series expansions is employed in this report to obtain approximations to the test statistics of some optimal detectors. The Edgeworth series is used to devise an approximate locally optimal detection scheme. Convergence properties are addressed here. Under a certain condition, the Tauberian condition, this approximation converges to the locally optimal statistic. However, it is shown that, even if this condition is not satisfied, the approximate scheme may still provide "reasonably good" performance, e.g. the ARE being greater than unity as compared to the linear detector. The other achievement of this study

in the application of series expansions is an investigation of the effect of the noise skewness on the detector performance. Asymptotic relations between the performance and the noise skewness are obtained for both the sample mean detector and the sign detector. The sign detector is shown to be less sensitive than the sample mean detector to the noise skewness, an intuitively reasonable result. Furthermore, a modification of the sample mean detector is proposed. Performance of this modified scheme in skewed noise is asymptotically equivalent to that of the sample mean detector in Gaussian noise.

The second part of this report is concerned with signal detection in bounded noise. When the signal is assumed to be a known constant, a sequential solution is obtained. By assuming that the bound is a random variable, a random test results. Evaluation of this test thus

requires the statistical moments of the false-alarm rate and power. On the other hand, if the signal is an unknown constant, an estimation-detection procedure is proposed. Unfortunately, performance evaluation of the procedure is still an unresolved question. Finally, a set-theoretic formulation of a vector-signal detection problem is discussed. Again, estimates of the signal are obtained to make decisions as to whether or not the signal is present. Since no statistical assumptions are imposed here, performance can not be evaluated conventionally.

6.2 Future Research

The problem of modifying the sample mean detector (or the linear detector) using methods of series expansions has been addressed here. More sophisticated series expansions may be considered to achieve better approximations and convergence results. As for examining the noise skewness on detector performance, the analysis here is based on large sample size and small skewness assumptions. Other methods may be used to remove these two assumptions.

The performance evaluation for signal detection in bounded noise is not yet well-resolved. Furthermore, the number of samples required to terminate the sequential procedure may be large. Some revision of this scheme will be an interesting problem. Finally, it is the author's belief that the set-theoretic formulations discussed in Chapter 5 deserves more consideration. Failure detection of dynamical systems may be one of the applications.

OFFICE OF NAVAL RESEARCH
STATISTICS AND PROBABILITY PROGRAM

BASIC DISTRIBUTION LIST
FOR
UNCLASSIFIED TECHNICAL REPORTS

FEBRUARY   1982

Copies   Copies

Statistics and Probability   Navy Library
 Program (Code 411(SP))   National Space Technology Laboratory
Office of Naval Research   Attn: Navy Librarian
Arlington, VA   22217   3   Bay St. Louis, MS   39522   1

Defense Technical Information   U. S. Army Research Office
 Center   P.O. Box 12211
Cameron Station   Attn: Dr. J. Chandra
Alexandria, VA   22314   12   Research Triangle Park, NC
                                        27706   1

Commanding Officer
Office of Naval Research   Director
 Eastern/Central Regional Office   National Security Agency
Attn: Director for Science   Attn: R51, Dr. Maar
Barnes Building   Fort Meade, MD 20755   1
495 Summer Street
Boston, MA   02210   1   ATAA-SL, Library
                                      U.S. Army TRADOC Systems
                                       Analysis Activity
Commanding Officer   Department of the Army
Office of Naval Research   White Sands Missile Range, NM
 Western Regional Office   88002   1
Attn: Dr. Richard Lau
1030 East Green Street   ARI Field Unit-USAREUR
Pasadena, CA 91101   1   Attn: Library
                                      c/o ODCSPER
U. S. ONR Liaison Office - Far East   HQ USAEREUR & 7th Army
Attn: Scientific Director   APO New York   09403   1
APO San Francisco 96503   1

Applied Mathematics Laboratory   Library, Code 1424
David Taylor Naval Ship Research   Naval Postgraduate School
 and Development Center   Monterey, CA 93940   1
Attn: Mr. G. H. Gleissner
Bethesda, Maryland 20084   1   Technical Information Division
                                      Naval Research Laboratory
Commandant of the Marine Coprs   Washington, DC   20375   1
 (Code AX)
Attn: Dr. A. L. Slafkosky   OASD (I&L), Pentagon
       Scientific Advisor   Attn: Mr. Charles S. Smith
Washington, DC   20380   1   Washington, DC 20301   1

Copies

Copies

Director
AMSAA
Attn: DRXSY-MP, H. Cohen
Aberdeen Proving Ground, MD   1
  21005

Dr. Gerhard Heiche
Naval Air Systems Command
  (NAIR 03)
Jefferson Plaza No. 1
Arlington, VA   20360   1

Dr. Barbara Bailar
Associate Director, Statistical
  Standards
Bureau of Census
Washington, DC   20233   1

Leon Slavin
Naval Sea Systems Command
  (NSEA 05H)
Crystal Mall #4, Rm. 129
Washington, DC   20036   1

B. E. Clark
RR #2, Box 647-B
Graham, NC   27253   1

Naval Underwater Systems Center
Attn: Dr. Derrill J. Bordelon
      Code 601
Newport, Rhode Island   02840 1

Naval Coastal Systems Center
Code 741
Attn: Mr. C. M. Bennett
Panama City, FL   32401   1

Naval Electronic Systems Command
  (NELEX 612)
Attn: John Schuster
National Center No. 1
Arlington, VA   20360   1

Defense Logistics Studies
  Information Exchange
Army Logistics Management Center
Attn: Mr. J. Dowling
Fort Lee, VA   23801   1

Reliability Analysis Center (RAC)
RADC/RBRAC
Attn: I. L. Krulac
      Data Coordinator/
      Government Programs
Griffiss AFB, New York   13441   1

Technical Library
Naval Ordnance Station
Indian Head, MD   20640   1

Library
Naval Ocean Systems Center
San Diego, CA   92152   1

Technical Library
Bureau of Naval Personnel
Department of the Navy
Washington, DC   20370   1

Mr. Dan Leonard
Code 8105
Naval Ocean Systems Center
San Diego, CA   92152   1

Dr. Alan F. Petty
Code 7930
Naval Research Laboratory
Washington, DC   20375   1

Dr. M. J. Fischer
Defense Communications Agency
Defense Communications Engineering
  Center
1860 Wiehle Avenue
Reston, VA   22090   1

Mr. Jim Gates
Code 9211
Fleet Material Support Office
U. S. Navy Supply Center
Mechanicsburg, PA   17055   1

Mr. Ted Tupper
Code M-311C
Military Sealift Command
Department of the Navy
Washington, DC   20390   1

Mr. F. R. Del Priori
Code 224
Operational Test and Evaluation
 Force (OPTEVFOR)
Norfolk, VA 23511   1